THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

Vol. XLVIII APRIL, 1936 No. 2

NEURAL RESOURCES OF PSYCHOLOGICAL PERFORMANCE

By T. A. RYAN, Cornell University

After reviewing current and recent textbooks of psychology, Lashley once observed that "the chapter on the nervous system seems to provide an excuse for pictures in an otherwise dry and monotonous text. That it has any other function is not clear; there may be cursory references to it in the later chapters on instinct and habit, but where the problems of psychology become more complex and interesting, the nervous system is dispensed with."[1]

On the face of it, the above statement is painfully true, although this frank critic seems to have overlooked one important motive for including in these works a chapter on the nervous system. The chapter undoubtedly serves as a confession of faith in neurology, a confession that psychology may and must ultimately look to the nervous system for an understanding of performance and conduct.

We may admit, however, that many psychologists who write in this spirit of blind faith make very little attempt to discover how far neurological knowledge has actually progressed toward the desired goal. Usually the neurological chapter in our general psychologies includes a very general description of the structures of the nervous system, a description which might have been written in 1900 as well as today. As Lashley contends, the author usually stops here, and henceforth makes use of the nervous system only as a *deus ex machina* which he calls in only when he needs a pat explanation in terms of instinct, learning or habit.

On the other hand, there are authors who make use of the nervous system in quite a different way. These accept as an established fact some one of the common simple hypotheses concerning neural function. Then they

* Accepted for publication December 20, 1934. This study was directed by Professor Madison Bentley.

[1] K. S. Lashley, Basic neural mechanisms in behavior, *Psychol. Rev.*, 37, 1930, 1.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

try to fit all psychological facts which interest them into this schema. This is ordinarily accomplished by leaving many facts entirely out of account, and distorting others into a verbal congruity with the chosen hypothesis.

It is the purpose of the present paper to examine some of these popular hypotheses in the light of what is known about psychological performance, and then to inspect certain recent investigations in the field of neural function. From these investigations we propose to pick out facts which purport to throw light on psychological performances and to regard them from a psychological point of view.

I

The various theories of cerebral function current today among neurologists and psychologists are extremely conflicting and confused. They are conflicting because the available facts are many and diverse, depending heavily upon the particular method used in obtaining them; a method usually inadequate to the intricate problems which it is devised to solve. The theories are confused because the interpretation of the facts rests in part upon the notion which the neurologist has of the psychological performances of the organism, and preconceptions of what should happen in the cerebrum in carrying on these performances.

In formulating a theory of cortical function, the neurologist has derived factual material from human and comparative anatomy of the nervous system, from neuro-physiology, from clinical neurology, and from studies involving a correlation between animal behavior and neural development or experimental injury. But he has been influenced in his formulation not only by experimental results themselves, but by general modes of approaching problems and by other outside considerations which are not quite so apparent.

Thus the study of action currents in the peripheral nerve fiber led to the very convenient conception of the central nervous system as a net-work of conducting elements analogous to electric wires. As Herrick says,[2] "The neurologist's problem is somewhat like that of an electrician who is sent to a large factory building, of which he has no previous knowledge, with orders to prepare a complete wiring chart of the building." Herrick then goes on to describe the electrician as finding sources of current, as determining the kinds of current supplied, and as tracing leads to various lamps, motors and transformers. "So the neurologist builds up his wiring chart of the body and thereby secures tolerably reliable information about the

[2] C. J. Herrick, *Brains of Rats and Men*, 1926, 149.

functions of those deeper parts of the brain whose experimental study is very difficult."

Herrick forgets, however, one fundamental difference between the two situations, a difference which may not be entirely one of complexity. The electrician approaches his task with a background of knowledge about the properties of generators, transformers, and lamps, as well as about the conduction of current in a single wire. Moreover, he knows what may be expected to happen to the electrical current when it passes such a complicated apparatus as a transformer. The neurologist, on the other hand, has some factual knowledge of the properties of peripheral nerve fibers, such as latent period, speed of conduction, action current, and of the receptor and effector organs to which they are connected, and he has to infer from these the functional properties and conditions of performance of the neurons in the brain. Because of the tempting analogy which we find in peripheral nerves to physical electrical systems, the neurologist has been led to assume that, throughout the nervous system, function is to be regarded in terms of connection from one neuron to the next, and in terms of changes in conductivity according to very simple principles.

This tracing out of connections is, of course, a necessary step toward gaining a knowledge of the nervous system, but it has been too often assumed that such knowledge when complete would tell most of the story. For example, the neurologist often states that the 'function' of a certain nucleus is visual, whereas all he knows is that the nucleus is connected to the optic tracts. No one has denied that the functions of the cortex are extremely complex, but the tendency has been to throw the complexity into terms of multiplicity of connection (which we might speak of as an arithmetical complexity), and to avoid discussion of the possible complexity of physiological conditions in the cortex.

The tendency toward simplification, however, came not only from the anatomist and the neurophysiologist, but has been strengthened from other sources as well. Learning, particularly animal learning, has often been regarded by experimenters in that field as primarily governed by bare repetition of the performance, and other conditions have been minimized. At least, investigators seem to forget these other conditions when they come to formulate a theoretical explanation of learning.

Thus we have a whole group of cortical theories based on the assumption that repetition is the essential factor in learning. Such theories have been carried along by many writers on learning, even though experimentation has shown that many other conditions are essential.

The theory that synaptic resistance is reduced by the passage of the

nerve impulse, and that the changes thus induced serve as the basis of plasticity of behavior; the hypothetical neural 'explanation' of the facts of conditioned reflexes; and Kappers' theory of the growth of cell-processes according to the principle of neurobiotaxis, are all widely accepted explanations which rest on the assumption of the primary importance of repetition. Even if we grant great importance to repetition, such conceptions have the further disadvantage of resting upon purely gratuitous views of neural conduction. When it was found that the synapse had the property of conducting the neural impulse in ony one direction, it was tempting to speculate on other possible properties which the synapse might possess. The proponents of the first of these theories hit on the possibility of variability of resistance as a convenient and plausible means of explaining learning, without having to abandon the electric-wire analogy in the nervous system. But, as Lashley says, "there is no evidence that synapses vary in resistance, or that, if they do, the resistance is altered by the passage of the nerve impulse."[3]

The second theory is in reality not an explanation at all, since it merely states that when two stimuli are presented together for a number of times, central connections undergo some change. For an explanation of these changes in connection it has to fall back on some of the other theories, like that of changes of resistance in the synapse, or venture no explanation at all. As regards the third theory, growth of the cell processes with the rapidity required for learning has not been demonstrated any more than has variable resistance of the synapse.[4]

It may be objected that we cannot criticize a theory on the ground that it has not been demonstrated to be true, but only on the ground that it does not cover all the facts which it is offered to explain. But there is evidence, both neurological and psychological, that we must look for changes in the central nervous system which are more complicated than those which these theories assume, in accounting for the effects of past on present performances. We shall take up later, in dealing with the work of Lashley, some of the neurological evidence on this point. For the present, it may be well to point out briefly certain general results of investigations on learning and memory, to show what an adequate cerebral theory must take into account, if such a theory is ever to be evolved.

The theories of learning mentioned above have been little more than restatements in physiological terms of the 'law' of association, which states

[3] K. S. Lashley, Basic neural mechanisms in behavior, *Psychol. Rev.*, 37, 1930, 2.
[4] Lashley, Nervous mechanisms in learning, *Foundations of Experimental Psychology*, 1929, 524-563.

that, if two items are repeatedly present together, the subsequent occurrence of one will tend to recall the other. The main difference is that these more modern theories, instead of postulating bonds which are built up between the items themselves, have assumed that some sort of functional readiness has been brought about in the cortex by repetition. These notions had their value in breaking down the old conceptions of the storing up of ideas or images, but we must look for others which will take account of the advances in psychological knowledge, if we wish to devise working hypotheses.

First, it must be recognized that learning is more than a tying together of successive muscular movements in the maze, or of groups of nonsense syllables, by dint of repetition alone. If we look at the learning which goes on in everyday activity, or even if we look more closely at the animal running the maze, or at ourselves learning a list of nonsense syllables, we see that what occurs is a general reorganization of the activity. Originally disparate items are ligated, it is true, but they are ligated in such a way that the result is a single total performance carried out under a single determination, rather than a chain of activities tied together. All the while this is going on, items originally present are eliminated, governing factors at first not operative are included in the situation or take the place of others which drop out.

Proponents of the theories of changes in synaptic resistance or of conditioned reflexes have tried to fit such things as substitution and inclusion into their schemas, and these may fit in if we grant their initial assumptions. But they have missed the fact that ligation is an organization into a total performance, a point which the writers on Gestalt theory have emphasized. Lewin,[5] for example, points out that an observer who has learned a series of syllables to the point where he can recite them as a whole, cannot give the next following syllable if he is presented with a syllable originally in the middle of the series. Also, elimination—which is an important change involved in learning—is scarcely taken account of at all in these theories.

In addition, and most importantly, it has been found that governing factors are more essential for learning than is mere repetition. The investigator of animal learning in the maze has regarded internal conditions merely as incentives for repetition on the part of the animal, and has forgotten that they may be as important for success as the repetition itself is. For obvious reasons, it is difficult to find out what are the internal con-

[5] Kurt Lewin, Das Problem der Willensmessung und das Grundgesetz der Assoziation, II, *Psychol. Forsch.*, 2, 1922, 121.

ditions or governors operative in animal learning. In experiments with man the case is clearer. Here we may induce an observer to repeat a performance without the instruction to learn, or with the suggestion that he will not be expected to learn, and it has been shown definitely that repetition may here become much less effective. The reason this fact was long overlooked is that, in most experiments, the observer not only repeats the material, but he repeats it with the instruction to learn. Not only that, but having learned a group of materials in such a situation, an observer will tend to proceed under the same instruction in similar situations later, even though he has not been given specific instructions 'to learn.' If we were to fit these facts into a theory of synaptic resistance, or of new pathways laid down by conditioning, we should be faced with the problem of explaining how instruction or task may facilitate or hinder changes in synaptic resistance, or the laying down of new pathways.

Memory proper, that is, the apprehension of objects and events as of the organism's own past, is something quite different in character from performing an activity which is 'learned.' In other words, remembering the person as having greeted you on the street yesterday is a different sort of performance from skilled activities like typing or reciting a well-known poem. Not as much experimental work has been done on the conditions of remembering, primarily because remembering has been regarded as just one form of learning, and it was assumed that the same conditions should hold true for both. It has been well established, however, that the most vivid memories are not those of items which have been repeated a number of times. In fact, repetition of perceiving sometimes dulls and blurs the remembered objects. Here, then, we find the most striking exception to the 'rule' that repetition is the most effective means of producing effects by past on present performances. Besides accounting for this fact, the neurologist must account for the fact that, in remembering, the organism is able to perform entirely by central means what it did before with specific aid of sense organs. It is probably correct to say that this is just one instance of the increasing cortical independence of extraorganic agencies, but the real problem is to find out how this independence comes about, and how the cortex operates when it is carrying on these more independent functions.[6]

So far we have dealt only with ways of accounting for plasticity of cortical functions, since the more specific theories of these functions have

[6] For a coherent account of the basal psychological performances and of the principle of psychological government, as touched upon in this paper, see M. Bentley, *The New Field of Psychology*, 1934, v-ix, 3-31, 73-80, 89-90, 117-118, 175-208, 380-383, 393-396.

been, in the main, attempts to account for learning. When we look for specific information or for attempted explanations of other activities, we find a surprising lack. Most of the instances of perceiving and acting, for example, come from the clinical neurologist. Because of the long tradition of cerebral localization from which he comes, and because of his interest in practical diagnostic aids, the clinical neurologist has been able to tell us in a very rough way what happens to these functions when lesions to certain parts of the cortex occur. But beyond telling us that injury to a certain part of the cortex may cause disturbance of a certain sort, he cannot go. He can only conjecture what part the particular region of the cortex played in the function involved, and he has often been led to a 'faculty' notion of psychology, with his faculties inferred from the various types of disturbance found in clinical cases. The clinical methods of the neurologist have often been extremely crude but in such recent studies as those by Gelb and Goldstein steps are being made toward a more accurate determination of psychological function after injury to the cortex.

Along with clinical studies of this sort, we might group a great many of the neurological experiments on lower animals. Most of the early 'functional' experiments with animals, and a great many even today, follow closely the techniques of the clinical neurologist with human patients. The experimenter takes out by operation a portion of the central nervous system, and then observes the animal's behavior for a time after the operation. Finally, he kills the animal and sections the brain or cord. From this sectioning he can describe the nuclei and tracts injured by the operation, and he attempts to correlate this injury with his gross descriptions of behavior. Because of the difficulties of technique, such results are usually based on a very few animals, and no account is taken of individual variations. The description of behavior consists of listing a few outstanding oddities of activity on the part of the animal.

Such a procedure is obviously best suited to a study of motor activities of the animal, but the authors of such studies sometimes attempt to draw conclusions concerning perceptive involvement in the activity, particularly involvement of the kinesthetic sort. The generalities drawn from this method are subject to the same limitations as those from the clinical methods.

This type of work is of aid in that it gives the psychologist a rough estimate of the part which the cortex plays in psychological functions, and indicates parts of the cortex which he must not overlook, if he is to discover underlying cortical resources of a certain psychological performance.

How the cortex plays its various and complicated rôles awaits an answer which may come by the perfection of newer techniques of studying brain

metabolism, action currents, and brain chemistry. But if physiologists continue to work in isolation from psychology, and if psychologists continue to borrow theories from the neurologist without question, even these latter methods will be able to accomplish little more than the older ones. One cannot expect to find out the nature of the bodily support of human performances without knowing something of the nature of these performances and the conditions which guide them.

The psychologist has too meekly accepted neurological theories, and has even tried to warp his psychological facts into line with these theories, poorly supported as they are. Instead, the neurologist should find the psychologist ready to criticize actively in the light of established psychological facts, and ready to supply these facts as a necessary background for neurological research.

II

The task of writing an adequate account of the available knowledge of the bodily resources of psychological performance, and of evaluating it from a psychological point of view, would involve a detailed treatment of a vast body of literature. It would be a task far beyond the scope of this paper.

On the other hand, general reviews do not strike deeply enough into the roots of the problem to accomplish our purpose. In order to re-analyze the material from a psychological point of view, we must go back to basic and tacit assumptions and to experimental techniques. It is only in this way that we may find out just what functions are involved in a certain experiment, and so avoid the difficulties of description in ambiguous and poorly defined terms.

Such a complete critical valuation must eventually be carried out, especially as the evidence becomes more complete. For the purposes of the present paper, however, we shall deal only with two examples. The work on this general problem might be divided into two gross types, experimental and speculative. The former attempts to answer a specific question by specific means, the latter argues logically from more remote evidence to a general view of neural function.

As an example of careful investigations of the first sort, we have chosen the experiments of Lashley, all of which fit together as a unit, and stand as a mass of evidence against the traditional theories of learning. From the second group, we have chosen two books of Herrick, who has made important attempts to bring the results of comparative neurology into line with recent physiological conceptions, particularly those of Child.

The investigations of Lashley depart from the earlier methods of study

of the cerebral functions, in that we find there an application of methods developed in the psychological laboratory for animals to the study of cerebral lesions, a very careful technique for determining the exact locus and extent of the lesions produced, and statistical treatment of results obtained from a large number of animals. By this method, Lashley is able to avoid many of the pitfalls of casual observation of individual animals, and to disregard inconsistencies due to chance errors made in determining the nature of the lesions. The method also allows the animals to be grouped according to a number of criteria, so that the relative effects of different variables may be studied.

The introduction of a quantitative method into this field, although it may not give us the final answer to many questions, is nevertheless of value beyond the specific results which it produces. Very often those who merely observe the behavior of animals with cerebral lesions have regarded the functions of the organism as faculties which are either present or not present. Using such quantitative methods, the experimenter may see instead that these performances are not present or absent as units, but that they may be more or less disturbed under various conditions.

As for the results and interpretations which we find in Lashley's papers, we must remember that for an adequate criticism it is necessary to regard the course of his investigations as a whole, rather than by individual papers. Lashley's point of view today represents a long, gradual development, and shows an emphasis quite different from his point of view in, say, 1929. Each experiment leads to further work, which quite often changes the significance of the earlier data. Too often critics have regarded Lashley's statements as final, rather than as summing up the trend of his work to date.

History. This series of studies had its inception about 1917, when Lashley and Franz published a paper concerning the effects of frontal lesions upon the rat's learning of a simple maze. They found that rats with frontal lesions can learn a simple maze as well as normal animals. Lashley continued the investigation by means of extended experiments with learning in various kinds of problem and with various kinds of lesion. He found that the double-platform-box problem is learned as quickly by animals with frontal lesions as it is by normal animals. In fact, the operated animals seem to be slightly more efficient.

The next part of the series consisted of a group of studies in brightness discrimination lasting until 1926. Here Lashley made a surprising discovery.[7] Rats can learn to go to a bright box to obtain food and avoid a dark box just as well with all or part of the occipital cortex removed as other animals who have undergone no experimental injuries. If the operation is performed *after* the animal has been trained, however, the occipital lesions do disturb subsequent performance.

[7] K. S. Lashley, Studies in cerebral function in learning: VII. The relation between cerebral mass, learning, and retention, *J. Comp. Neurol.*, 41, 1926, 1-58.

Under these conditions, the loss in efficiency of performance after the operation is roughly proportional to the total amount of occipital cortex destroyed, irrespective of the part of this region involved; but injuries to other regions of the cortex have no effect on the performance.

The application of this technique to more complicated mazes led to a still more surprising result. In an extended investigation of mazes of varying complexity,[8] Lashley found that the learning of a simple maze of one cul-de-sac is only slightly retarded by lesions up to one-third of the cortex, and post-operative retention of the maze is slightly reduced. But with more complicated mazes (*e.g.* eight culs-de-sac) the correlation between the amount of cortical destruction and efficiency in learning becomes higher, and there is some correlation between the amount of lesion and the rate of 'forgetting.' 'Retention' of a maze learned before the operation is also reduced in a degree proportionate to the lesion. The most unexpected finding was that the loss did not seem to be significantly different for different regions of the cortex.

In the same volume, Lashley also summarizes his data on brightness discrimination, reported above, and the work on the platform-box by Lashley and Franz, Lashley, and Hunter. These latter studies support the original finding that this type of performance is not influenced by lesions up to 50% of the cortex, and is independent of the magnitude and locus of the lesion. They also show that retention of the performance is disturbed by lesions in the frontal third of the cortex, but not by lesions in any other region.

The differences found between the maze, on the one hand, and such problems as that of brightness discrimination, on the other, might be explained by assuming that the rat makes use of more sensory modalities in the maze than it does in the other situations, hence the equipotentiality of the cortex. In other words, the explanation might be in terms of direct interruption of projection fibers to the cortex, rather than in terms of cortical mechanisms themselves. But Lashley found that enucleation of the eyes does not produce noticeable defects in running a maze,[9] even though the animals have learned the maze before the enucleation is performed. When a previously blinded animal learns a maze, however, and then suffers destruction of all or part of the occipital cortex, retention is now disturbed in proportion to the amount of tissue destroyed. The occipital cortex must be involved, even though the animal does not make use of vision in learning.

Lashley and Ball found that maze performances are retained in animals which have had the dorsal or lateral funiculi of the spinal cord severed in the cervical region.[10] The animals showed disturbances of posture which would indicate that kinesthetic paths were interrupted, but they showed surprisingly little disturbance of motor coördination except for awkwardness of posture and difficulties in making the turns of the maze, in spite of the fact that both ascending and descending paths were cut. The results in maze learning are to be expected from Lashley's earlier discovery that extirpation of the stimulable cortex of the precentral region does not disturb retention of motor habits either with rats or with monkeys. For a part of Lashley and Ball's experiments, the mazes were darkened so that the results could

---

[8] Lashley, *Brain Mechanisms and Intelligence*, 1929, 1-186.
[9] *Brain Mechanisms and Intelligence*, 112.
[10] K. S. Lashley and J. Ball, Spinal conduction and kinesthetic sensitivity in the maze habit, *J. Comp. Psychol.*, 9, 1929, 71-106.

not depend upon a transfer from kinesthetic to visual cues. Olfactory and tactual cues were, Lashley believes, eliminated in all these investigations.

Such results were so inconsistent with the beautifully simple conceptions one would derive from the conventional diagrams of the 'conditioning process,' or any process of laying down specific pathways through synaptic changes or dendritic growth, that Lashley was led to a violent swing away from all conventional views. He concluded that an animal in running a maze, or in any other performance, makes use of any sensory cues available, and that such a performance involves nothing like the chain reflex postulated by Watson and later writers. The experiments of Lashley and McCarthy,[11] and Lashley's work on the precentral gyrus of monkeys, indicate that animals can retain a maze performance or manipulative movements in spite of motor difficulties which necessitate changes in the specific movements involved. In other words, an animal may carry out a learned performance with neural pathways and muscles not used in the original learning. To quote Lashley and Ball,[12] "The engram of the maze-habit consists of some central organization in which the general direction and succession of turns are so recorded that, once the series is initiated, the essential sequence of movements may be performed in the absence of sensory control and with considerable variation in the actual movements produced. This is, of course, a very vague statement and implies conceptions of plasticity in nervous functioning which run counter to the whole doctrine of conditioned reflexes, yet the problem seems at present to admit of no greater simplification."

From this point of view, all we can say about the part of the cortex in learning is that learning in some situations seems to involve the whole cortex equally, in others it involves only a portion of the cortex, but within this area no part seems to be more essential than another. What is more, the cortex may be necessary for the retention of a performance learned with the cortex intact, whereas destruction of the same area has no effect on efficiency of learning that same performance for the first time.

Lashley argues that if we are to hold to theories of conditioned reflex pathways, we must assume that the changes involved occur in large numbers in whole areas of the cortex, as Pavlov was forced to do.[13] But even if we were to assume this multiplication of conditioned arcs, we should still meet with difficulty. In some of the more complicated mazes used by Lashley, animals with a portion of the cortex removed learned more slowly than normal animals, it is true, but were able finally to reach Lashley's criteria of learning. After an interval of time, however, the operated animals are less efficient in running the maze than normal rats who have had the same period without practice. If, in both groups of animals, learning came about by effecting changes in functional synapses, why should the changes last a shorter time for animals lacking part of the cortex? We might go on and assume other conditions which are effective in changing synaptic connections, but as Lashley says, "the truth is that our hypotheses have so far exceeded our facts as to be beyond any experimental test, and the piling on of a few more speculations is likely to make the whole structure collapse of its own weight."[14]

---

[11] K. S. Lashley and D. A. McCarthy, The survival of the maze habit after cerebellar lesions, *J. Comp. Psychol.*, 6, 1926, 423-435.

[12] Lashley and Ball, *op. cit.*, 97.

[13] I. P. Pavlov, *Conditioned Reflexes*, 1927, 339.

[14] *Brain Mechanisms and Intelligence*, 127.

A second difficulty is that lesions which destroy large numbers of projection fibers without very large destruction of cortical tissue do not produce any more disturbance than other lesions of about the same size without involvement of association and projection fibers. In fact, "the destruction of the areas supplied by the projection fibers produces much more serious disturbance than an extensive injury to the fibers."[15] And again, in summary, "Contribution to the efficiency of performance of a habit by the summated action of equivalent bonds is understandable for habits which have already been formed. It is less obvious how their formation could be retarded by a reduction of the number of possible bonds before the bonds have been formed. For this to occur would require that only a limited number of bonds for any given performance can be formed readily in any part of the cortex—an assumption which is contradicted by the results with the visual discrimination habit which normally involves only the occipital cortex, yet is formed with equal facility in the absence of this area."[16]

Since this general summing-up in 1929, Lashley's work has taken two main directions, one involving further investigations with mazes, mainly to furnish further data for a more accurate statistical analysis, and the other a continuation of the series on the mechanism of vision.

The work in the former group represents little change from the material of *Brain Mechanisms and Intelligence,* except for matters of detail. Lashley and Wiley report an extended study of maze learning,[17] with a detailed analysis of results from a large number of rats. The previous finding that retardation of learning is highly correlated with the extent of lesions is corroborated, and the relation between the two is found to be curvilinear. But when the animals are divided into groups according to the cerebral area predominantly injured, the exact equipotentiality of the various regions previously predicted is not established. Although lesions anywhere in the cortex produce significant retardation in learning, the correlations are higher in the motor, visual, and somesthetic areas than they are in the auditory area.

Lashley and Wiley also found that, although the correlation between the performance in different mazes for the same normal animals is very low, there is a high correlation for the operated animals. In other words, individual animals varied widely in efficiency in mazes of varying complexity, whereas among the operated animals the individual animal which ran one maze poorly was also likely to run any of the others poorly.

In 1929 Lashley had suggested that cerebral lesions resulted in increase in difficulty of more complicated mazes relative to the more simple. This later study does not bear out this suggestion. With the type of mazes used here, the relative difficulty of the mazes for the operated animals is about the same as the relative difficulty for the normal rats. Lashley had also suggested that a possible interpretation of the earlier results was that small lesions have very little effect and that there is a critical amount of tissue beyond which the lesion is effective. The correlations alone do not show this, but in this investigation animals were used in sufficient numbers to permit dividing them into groups according to the degree of injury. The authors

[15] *Ibid.,* 130. [16] *Ibid.,* 131.

[17] K. S. Lashley and L. E. Wiley, Studies: IX. Mass action in relation to the number of elements in the problem to be learned, *J. Comp. Neurol.*, 57, 1933, 3-56.

conclude from this analysis that the effect of cortical lesions is continuous and shows no sudden jumps.

As Lashley and Wiley point out, the difficulty of deriving any positive conclusions concerning the nature of cerebral function from these results lies in the fact that the results are in terms of group-averages. Inspection of the individual cases shows that animals with similar amounts of lesion may differ widely in efficiency of running the mazes. They do not believe that there is sufficient individual variation in the anatomical structure of the cortex to account for this fact, nor do they believe that it can be explained in terms of chance success in solving the problems, unless the discovery of some general principle of maze-running is assumed. The possibility of individual differences in the sensory modalities made use of in the performance seems to be ruled out by the lack of effects on such performance of blinding and anesthesias of various sorts.

The authors suggest as an alternative that, since young animals and animals which are more efficient to start with (Lashley and Wiley say 'intelligent' animals), apparently show smaller effects from cerebral lesions, "severity of symptoms is dependent not only upon the locus and extent of lesion, but also upon the general level of dynamic functioning of the organism."[18] The conflicting nature of the evidence which we find concerning localization of function in the human nervous system might be attributed in part to similar individual variation in the effects of injury.

Such results seem to indicate that Lashley's method has yielded most of the important results which may be expected from its application to the learning of mazes. By devising means of application of the same technique to activities of a different sort, however, he has been able to produce quite different results. In the earlier studies, he had trained animals to discriminate between a lighted and a dark compartment, but in 1930 he applied the same method to discriminations involving smaller differences of intensity of illumination.[19] With this change in the situation, he found that lesions to the occipital cortex did cause retardation in learning the performance for the first time, and found a significant correlation between extent of lesion and retardation. Also the thresholds for discrimination were higher for the injured rats than they were for the normal animals.

Then, by devising an apparatus in which the rats jumped at cards bearing different geometrical designs in order to reach food, Lashley showed that rats could learn to discriminate between inverted and upright triangles, circles and triangles, and a number of other designs. The effects of cerebral lesions on this performance were found to be quite different from the effects on brightness discrimination.[20] Interruption of the optic radiations led to disturbances under which animals could not learn to distinguish designs or perceive the distance of objects, although the same animals could still learn to discriminate brightness. Destruction of a certain small area in the lateral part of the occipital region resulted in failure of the rats to learn to discriminate any of the geometric patterns used, but the animals could

---

[18] Lashley and Wiley, *op. cit.*, 47.

[19] Lashley, The mechanism of vision: II. The influence of cerebral lesions upon the threshold of discrimination for brightness in the rat, *Ped. Sem.*, 37, 1930, 461-480.

[20] Lashley, The mechanism of vision: IV. The cerebral areas necessary for pattern vision in the rat, *J. Comp. Neurol.*, 53, 1931, 419-478.

still distinguish direction and to some extent distance of objects to which they were to jump. Lesions in the supposed visual association area, or in any other part of the cortex except the small area already mentioned, had no effect on any of these performances.

A second analysis[21] of the data on brightness discrimination in the light of this discovery brought a drastic revision of the conclusions formerly drawn.[22] On re-examination of the 1926 data, Lashley found that lesions which had not invaded this small area critical for pattern vision had not produced any postoperative amnesia for the activity of brightness-discrimination. On the other hand, lesions which had totally destroyed this area but had involved little other tissue had not entailed total loss of retention. The total degree of amnesia was proportional to the amount of tissue destroyed *beyond* the critical area. In other words, as long as the critical area was destroyed, the amount of disturbance was proportional to the size of the lesion. Lashley does not believe that this disturbance is due to the loss of pattern-vision, and believes that the animals had not been discriminating on the basis of differences of (spatial) pattern in the original learning.

In order to determine what relation this small critical region bears to the projection of the visual fibers in the cortex, Lashley had next to determine the course of the visual fibers from the retina to the lateral geniculate bodies, and of the projection fibers from there to the cortex. He found by the method of degeneration that there is a rough spatial representation of the retina, both in the geniculate bodies and in the cortex.

With this knowledge as a background, further study of learning to discriminate patterns, and the effects of lesions on retention of these performances, was carried on by Lashley and Frank.[23] The authors found that the previously learned performance could still be carried out as long as part of the projection area for the temporal retina of one side was left uninjured. This Lashley believes to be the part of the retinal field used by the rat in binocular vision, because of its limitations of convergence. He also believes that the results are due to blindness of the operated rats for geometrical designs, since degeneration occurs in the lateral geniculate bodies after cortical lesions in the projection areas. In other words, he attributes his results to direct consequences in visual perceiving, rather than to anything like 'visual amnesia.'

Discussion and interpretation. Lashley's attack on the explanations of maze running in terms of specific sensory cues, in particular kinesthetic cues, was an approach from the physiological side to a characteristic of learning which we have discussed from a psychological point of view in the introduction to this paper. We have said that as a performance is being automatized, one thing which occurs is an organization into a single total

[21] Lashley, Studies of cerebral function in learning: VIII. A reanalysis of data on mass action in the visual cortex, *ibid.*, 54, 1932, 77-84.

[22] Lashley, Studies of cerebral function in learning: VII. The relation between cerebral mass, learning, and retention, *ibid.*, 41, 1926, 1-58.

[23] K. S. Lashley and M. Frank, The mechanism of vision: X. Post-operative disturbance of habits based on detail vision in the rat after lesions in the cerebral visual areas, *J. Comp. Psychol.*, 17, 1934, 355-392.

performance which runs its course under a single central determination, a ligation of originally disparate movements which is something more than a mere tying together of successive items by means of 'bonds' of some sort. So much has become established by means of psychological investigations of learning.

In their paper dealing with interruption of spinal pathways and learning, Lashley and Ball reached a conclusion concerning the neural resources in maze learning which must lead essentially to the same point of view. We have already quoted briefly from the paper in our history of Lashley's work, but it may be well to emphasize their conclusions again in this connection.

After discussing the various possible explanations for their results, the authors suggest that "the engram of the maze habit consists of some central organization in which the general direction and succession of turns are so recorded that, once the series is initiated, the essential sequence of movements may be performed in the absence of sensory control and with considerable variation in the actual movements produced."[24] Again, "The results of these studies seem explicable only on the assumption that the speed and extent of movement (intensity and duration of muscular innervation) may be determined in the initial set before any overt movement has actually occurred, and that 'current control' enters in only when there is some departure from the conditions for which the original adjustment was made."[25]

Lashley admits that such statements are vague and radically different from those generally current, but he believes that they have more justification in fact than more concrete and deceivingly 'factual' theories of 'chain reflexes.' We may point out that it is a statement which is closer, as well, to the facts of learning as determined by psychological investigation.

It is unfortunate, however, that Lashley did not state more definitely what he meant by 'sensory control' in the above quotations. When he says that a rat may run a maze 'without sensory control,' it seems evident from the context—and as he later confirms[26]—that he means that the initiation of each succeeding muscle-contraction does not depend upon kinesthetic stimulation from a certain specific muscle-contraction preceding it. In other words, the animal running a maze is directed primarily by conditions in the central nervous system which are brought about by previous performance in the maze. But any sensory stimulations available may be made use of as more general and subsidiary aids in directing the performance, rather than as specific initiating factors in the individual movements.

It is, of course, inconceivable how a rat, no matter how highly automatized, could run a maze without making use of any of its sense organs.

[24] Lashley and Ball, *op. cit.*, 97. [25] *Ibid.*, 99.
[26] Lashley, Cerebral control versus reflexology: A reply to Professor Hunter, *J. Gen. Psychol.*, 5, 1931, 3-20.

This latter interpretation is, however, apparently the one taken by Hunter, and Honzik, the former in criticizing, and the latter in supporting Lashley's statements.

Hunter performed experiments in which rats learned mazes in which certain sensory cues were made available.[27] Then, when these cues were removed or mixed up in some way, the rats made errors in the maze. This is, of course, inconsistent with Lashley's conclusion only if we accept this interpretation.

Hunter's difficulty is, however, more deep-rooted than a mere misinterpretation of a single statement. The following quotation from his article bears this out. "The explanation of the double-alternation-temporal maze habit thus requires the assumption that the stimuli involved in running such a maze are supplemented by some symbolic process, or by some central neural engram. Of the two possibilities, I favor the first, partly because such a factor is known to exist in some animals, and partly because it seems scientifically sounder to exhaust the peripheral possibilities of explanation before resorting to central possibilities."[28]

How such a 'symbolic process,' whatever that may mean, can be accounted for in terms of peripheral processes without an appeal to central processes, Hunter does not explain.

Hunter goes on to say that instruction may exert control over a whole series of activities, even though the instruction is given only once. But he claims that it is unnecessary to appeal to anything like 'cortical set,' since an observer may repeat the words of the instructions at intervals during the series of activities. Apparently Hunter does not regard an explanation in neural terms of the effects of instruction as any more difficult than an explanation or description in neural terms of the knee-jerk. It is the common behavioristic assumption that a description of all the stimuli involved in an activity can tell us the whole story, an assumption which is used as an excuse to avoid discussion of the intricate activities of the central nervous system.

Honzik leans to the other extreme in his experiments purporting to support Lashley's conclusions, taking the phrase 'without sensory control' literally. Honzik's rats were blinded, and then trained on elevated mazes in which the various sections were interchanged, the whole maze changed in its orientation at various times, while the food-box, where the animals were fed was sometimes present at the end of the maze and sometimes removed. The only item which remained constant was the pattern of the maze. Since the rats were able to learn to run the maze without error under these conditions, Honzik believes that he is justified in assuming that the animals learned "independently of specific visual, tactual, olfactory, and auditory cues, or stimuli."[29]

When the rats had learned to run the maze without error under the above conditions, Honzik introduced a new variation. At various points in the middle of the maze he inserted short-cuts, so that the rats had to run the maze with one section which had been present in the original training now omitted. He found that, when

[27] W. S. Hunter, A further consideration of the sensory control of the maze habit in the white rat, *Ped. Sem.*, 38, 1930, 3-19.

[28] Hunter, *op. cit.*, 3.

[29] C. H. Honzik, Cerebral control in the maze learning of rats, *J. Comp. Psychol.*, 15, 1933, 96.

the short-cut was introduced, the rats made only slightly poorer scores than normal. If the rat were put in the middle of the maze, however, he made significantly more errors. The results were by no means uniform, and the ease with which rats were able to take such short-cuts seemed to depend to some extent upon the position and general orientation of the short-cuts. Honzik, however, minimizes these differences, and treats his results as uniform in drawing his conclusions. He writes, "although a part of the maze was cut out, the run to the cut out portion plus the run over the gap enabled the rat to 'know' its whereabouts. And this knowledge of whereabouts, it is quite certain, was independent of specific directive sensory stimuli. How were the rats enabled to run with customary accuracy in the various short-cut trials? This paper has been entitled 'Cerebral Control in Maze Learning,' the implication being that the capacity in question is some function of the cerebrum."[20]

Deliberately to minimize peripheral conditions for learning, and to refer everything to a vague cerebral 'capacity,' as Honzik does, certainly offers little of value to the understanding of cerebral function. Honzik's results furnish interesting evidence against over-simplification in terms of chain-reflexes or S-R formulæ, but they do not mean that the cerebrum operates *in vacuo*. If he had studied his results with the purpose of finding out how it is that the rat 'knows' where he is after running across a short-cut, and under what conditions he does 'know,' we should have had a much more significant contribution to the problem. Apparently he assumes that if he were to find some external or peripheral conditions operating, it would cast a slur upon the importance of the cerebrum. This, as we have pointed out, is the attitude taken by Hunter, and many others as well, when they minimize the rôle of the central nervous system, pointing out that, since there are external conditions present, we do not need to speak of central control.

The concept of 'mass action' has played such an important rôle with Lashley's critics that it may be well to discuss it here in relation to Lashley's later studies. The theory has taken on an all-inclusive character, particularly for those who have not followed Lashley's investigations closely, a character which Lashley does not ascribe to it, and which is not supported by his experiments.

We quote Lashley's statement in 1931.

"The term *equipotentiality* was used to describe the condition in which lesion to any part of a cerebral area produced no measurable disturbance of function, whereas destruction of the entire area was followed by conspicuous loss of function. More recent data, indicating that cerebral lesion always produces some deterioration of all the functions of the injured area, have led to the more comprehensive notion of *mass action*. This may be defined briefly as the theory that, within any functional area of the cerebral cortex, the subordinate parts can individually perform any of the functions of the whole, although at a lower level of efficiency. I have insisted that the

[20] Honzik., *op. cit.*, 131.

theory is applicable only to certain functions which cannot yet be defined in general terms.[21]

At first glance, it might be assumed that Lashley's latest papers are directly contradictory to a theory of "mass action"; but it must be noted that he says, "I have insisted that the theory is applicable only to certain functions which cannot yet be defined in general terms." He has shown definitely that the visual cortex must be involved in different ways when a rat is engaged in discriminating dark from light, when it is discriminating unlike geometrical figures, and when a blind rat is learning to run a maze.

In the first of these activities, the visual cortex may be removed without increasing the number of errors made in learning the performance for the first time, but after the rat has learned the performance with the cortex intact, removal of the occipital cortex leads to marked increase in errors following the operation. When the animal is engaged in discriminating geometrical designs, the same area of the cortex is necessary both for learning and 'retention.' The visual cortex cannot be involved at all as a projection area when a blind rat is learning a maze, yet destruction of the visual area after the rat has learned the maze disturbs its later performance in the maze.

It is useless to point out, as Herrick does, that maze-running and brightness-discrimination are probably essentially subcortical functions. Since Herrick does admit that the cortex is involved in learning such performances, he is really saying nothing more than that the cortex is involved in a different way, and it is dangerous to classify the performances as essentially subcortical. We might also say that the visual projection areas are involved in a 'primary' way in the perception of geometrical designs, and in a 'secondary' way in these other activities, but this again would tell us little. Attaching such labels brings us but little nearer to the actual mechanisms involved.

Since Lashley has limited himself to a quantitative determination of changes in efficiency under cerebral insult, his results are valid only as far as these measures are concerned. If a certain lesion produces a marked increase in the number of errors in a maze, he knows that the lesion has produced some effect on that activity. What change has occurred, and how the change is brought about, is not a question which can be answered by his method. In Chapter IX of *Brain Mechanisms and Intelligence* Lashley reports some descriptions of the behavior of animals in the maze, but the differences which can be determined by such observation are too gross to tell us anything further about the effects of the lesions on performance.

Hamilton and Ellis have devised ingenious techniques for bringing out the differences in behavior between injured and normal rats, and they have analyzed the types of errors made in the two groups.[22] They have not, however, made any dis-

[21] K. S. Lashley, Cerebral control versus reflexology: A reply to Professor Hunter, *J. Gen. Psychol.*, 5, 1931, 4 (footnote 3).

[22] J. A. Hamilton and W. D. Ellis, Behavior constancy in rats, *Ped. Sem.*, 42, 1933, 120-138; Persistence and behavior constancy, *idem*, 140-151; Behavior constancy, *J. Gen. Psychol.*, 8, 1933, 421-429.

tinctions between operated animals with various types of lesion, nor have they brought out all the material which might be derived from their protocols. With a more extensive application of methods similar to theirs, and with more detailed scrutiny of the results than they have used, results might possibly be obtained which would give us a more detailed knowledge of the actual effects of these lesions on performance.

Lashley is definitely limited by his method and makes no attempt to go beyond his limitations. On the basis of his results he can only point out that the same region of the cortex may be involved in different ways in different activities. He makes no attempt to guess what these several mechanisms may be.

The fact that the results of these investigations can be expressed only in such general terms as these does not, however, make possible a simple view of cerebral function. Any hypothesis which would allow the same neuron and the same cortical region to take on different *kinds* of activity at different times must of necessity be more complex than any of the prevailing theories. These theories certainly cannot successfully cope with the large mass of material which Lashley has brought forward. In fact, probably one of Lashley's greatest contributions has been to show the limitations of such theories, as well as the limitations of our knowledge of the functions of the central nervous system at large.

In Herrick's more general writings on the development and functions of the nervous system, we find an excellent example of the mixed sources from which current notions of the nervous system derive. Herrick's own researches have been in the field of comparative neurology. To this background he has added the influence of Child's conception of physiological gradients, and, like most neurologists, the schematic notions of reflex mechanisms evolved by Sherrington and later neurophysiologists. On the other hand, even though one of his books is called *Neurological Foundations of Animal Behavior*,[33] most of the material concerning animal behavior has been derived from biological sources and consists in references to 'intelligence,' 'physiological' and 'associative' memory, 'stereotyped behavior,' all used in the loose sense in which biologists are wont to use these terms.

We see the influence of Child most prominently in *Neurological Foundations*, which we shall discuss first. Child's conception of physiological gradients and physiological dominance, which he has presented systematically in his *Physiological Foundations*,[34] furnishes a useful way of describing the way organisms develop, in terms of active interrelations between various regions. In experiments on lower

[33] C. J. Herrick, *Neurological Foundations of Animal Behavior*, 1924, 1-334.
[34] C. M. Child, *Physiological Foundations of Behavior*, 1924, 1-330.

animals and on plants, he found that each organism showed evidence that certain parts of the organism (*e.g.* in animals, the head-region) were carrying on more intensive metabolic activity than others. Regions of high metabolic rate exert dominance over the development of other parts, in the sense that changes which occur in the development of these highly active regions also affect the development of other, less active, parts. The course of development of any cell of the organism depends, then, not only upon the nature of the chromosomes which it contains, but also upon the position which it occupies in these physiological gradients. Since physiological gradients are subject to change under a multiplicity of external factors, we are able, by studying these factors, to discover many of the external governors of development.

In his *Neurological Foundations* Herrick has used this conception in interpreting the facts of development of the nervous system. Since these relations of dominance depend not only upon the relative metabolic rates, but also upon a mode of conduction of some sort from one part of the organism to another, the development in the organism of a nervous system as a structure of high conductivity and rapid metabolic change is of great importance for these physiological gradients. It makes for a more intimate dependence of one organ of the body upon another, and it also profoundly complicates the pattern of gradients.

We are chiefly concerned, however, with the significance of the cerebral cortex for psychological function, and we shall therefore proceed at once to discover what light Herrick's treatment throws on this topic. He postulates within the central nervous system two types of tissue, one 'mature' type which develops early in ontogenetic and phylogenetic history, and now makes up fixed structural formations and functional mechanisms. The other is a 'young' type of tissue which is more plastic, being more easily modified by the physiological forces about it. All correlation centers are supposed to be made of both of these types of tissue in various ratios, but the amount of fixed tissue is relatively greater in the brain stem and in lower centers than it is in the cerebral cortex. The young tissues, predominant in the cortex, carry on metabolic activities at a higher level than the more stable mature tissues, hence the cortex dominates the whole organism.

It may be questioned whether we have any factual evidence that genetically young neural tissues do carry on metabolic activities at a higher level than do other tissues. We must accept this if we are to accept the tenet that the young tissues of the cortex dominate other tissues, for Herrick assumes that this dominance is a function of metabolic rate and conduction facilities. The latter are certainly much the same for either the young or mature types of tissue. He points out that nerve tissues perform two types of metabolic activity, one slow process of building up tissue, and the other the rapid metabolism of conduction. Presumably the high metabolic activity of the young tissue is of the former type and it is this which determines its domi-

nance. He does not make it clear, however, what part the second type of metabolic activity plays in these gradients. As Child uses the concept for simpler animals, physiological gradients have some basis in experiment, and are simple and well-defined. It is, however, a long step from these to the metabolic patterns in the highly complex organisms possessing cerebral cortexes. However this may be, it is well to consider whether such a treatment has brought anything in the way of a fresh point of view, and if possible what it may bring in the future.

The advantage of such a treatment is that it takes the emphasis from the individual reflex arcs within the central nervous system and gives us some sort of picture of the functional relationships in the nervous system as a whole. The concept of cerebral dominance and control is by no means new, but it has always been described in terms of reflex arcs and indivdual motor units within the central nervous system. The physiologist who is dealing with digestive processes, must, to obtain a complete picture of the significance of liver secretions, know the structure of the liver cell, the nature of the chemical secretions which it turns out, and the nature of the processes which go on within it. He must also know how hepatic function is regulated by, and what part it plays in, the digestive functions at large. The neurologist has been able to find out something of the histological nature of neurons, their *structural* relations to each other and to other parts of the body, and he is beginning to find out something about the chemical processes which go on in the individual cells. But how these processes are regulated by what is going on in other parts of the nervous system and the organism as a whole, and what significance these regulations have for the functions of the central nervous system as a whole, are questions concerning which he has for the most part only dubious inferences drawn from histological tracing of conduction pathways and from clinical cases.

It is the first of these latter problems; namely, how neural processes are regulated by what is going on in other parts of the organism, for which Herrick brings in Child's conception of physiological dominance. Any attempt which even states this problem explicitly must be regarded as a step in the right direction, but Herrick has done little more than to state Child's notions in terms of organisms of greater complexity. For the rest of his foundations of behavior he reverts to conventional materials, comparative morphological studies, reflex mechanisms, and so on.

In deriving significance from comparative morphological facts, Herrick always states his problems in very gross terms. Birds, for example, have highly developed basal ganglia but no cerebral cortex, and the activities are limited to very stereotyped

forms. On the other hand, mammals have a relatively large cerebral cortex, and these animals are more capable of learning and memory. Therefore in mammals the corpus striatum, or basal ganglion, deals with the stereotyped modes of behavior, whereas the cortex deals with more variable modes. Such logic, which is by no means new, neglects entirely the vast reorganization which has taken place in the brain between the birds and higher mammals, and regards the corpus striatum as a unit which remains fixed and unchanged in function in the midst of all these changes, merely because its appearance under the microscope is fairly stable. There is, however, some evidence that the basal ganglia have a quite different significance in those animals which possess a cerebral cortex.

The point with which I wish to deal here is, however, not the soundness of such a theory, but its value. Granting that such a conclusion is justified, how significant is it? The piling up of such elaborate evidence to show that the cortex has something to do with learning and memory seems unnecessary, and is, in fact, deceiving. One might be led to believe that these conclusions throw some light upon the nature of the contribution of the cerebral cortex to psychological performance.

Just as long as the neurologist regards the problem of neurological bases of learning and memory as a unitary problem, as Herrick does, he will be balked by his very statement of the problem. As we have previously pointed out, remembering is something quite different, psychologically, from learning, and also learning is only a general term which covers a vast number of changes in performance under various conditions. What right have we to assume that all of these have a common bodily basis?

If we are to get the most out of our comparative neurological evidence, we must first of all have an intimate knowledge of the specific performances of which the various animals under consideration are capable. To say that birds are capable only of stereotyped activities, while the higher mammals are capable to a greater degree of memory and learning, scarcely implies an intimate knowledge of specific performances under a variety of conditions. Secondly, we must attempt to get a picture of the nervous system as a whole, rather than to split it into units, each of which is to be regarded separately. To return to our former example, we must not be misled by the apparent structural similarity of the corpus striatum into forgetting that its functional relations with the rest of the nervous system must be vastly changed by the development of the cerebral cortex. It is because neurologists have not done these two things that they have been able to offer so little to the psychologist seeking light on the bodily resources of performance.

The term 'plastic' Herrick, with many other biologists and neurologists, uses very often as an explanatory term, whereas it is really descriptive

of variable performance. Herrick uses it as descriptive of a certain type of nervous tissue which is predominant in the cortex. Referring to Herrick's statements concerning the cortex mentioned above, those animals which have the most widely variable modes of performance possess cerebral cortexes, therefore, the cortex must be involved in such changes in performance. Describing the cortex as 'plastic' is supposed to give us a notion *how* it plays a part in these changes. Applying such a label, however, tells us essentially no more than that the cortex is involved in variable performance. It tells us nothing of the mechanism operating.

Of course, Herrick does go on to deal in somewhat more detail with the hypothetical nature of this plasticity. In synaptic nervous systems there are supposed to be two types of plasticity in correlation centers. He writes:

"The first is the presence of a switchboard type of structure in the correlation centers which permits a considerable variety of possible responses to any stimulus depending upon the pre-established arrangement of the correlation tracts, etc., and the internal physiological condition (fatigue, practice, attention, etc.) of the system involved at the moment of its activation. The second form of plasticity is the capacity for enlargement of the switchboard, through irradiation of nervous currents, facilitation by use, etc., resulting in an increase in the number of things one can do in response to a given sensory excitation. This is education."[35]

Earlier in the volume, Herrick gives us a neat illustration of what he means by this 'switchboard' arrangement in the correlation centers. Here he draws a diagram of a gustatory nerve fiber and a tactile nerve fiber converging upon a single correlation neuron, which in turn connects with a motor neuron. The mechanism of correlation is conceived thus:

"If stimulated by something of possible food value simultaneously from both organs of touch and taste buds, the excitations may reinforce each other (if the food feels right and tastes right) or they may interfere by reason of antagonism between their normal or typical responses (if the food feels right but has an objectionable taste)."[36]

This he calls the "simplest possible arrangement of the sort," and one must admit that it *is* simple. That it is a possible arrangement, or even a conceivable one, is another question.

Herrick is of course justified in his anatomical picture as far as it goes, for there are connections of the sort to be found. But, even though the illustration is drawn from the larval salamander, he certainly has not given us a complete picture. To treat a tactile or gustatory nerve as a single fiber making connection with a single type of neuron is bad enough anatomically, but when we look at it functionally it is even worse. Herrick seems to think of "food tastes right" as a characteristic of a nerve impulse itself, or at least as a message which is transmitted along a nerve fiber.

Of course, Herrick would admit that this is false, just as any other com-

[35] Herrick, *Neurological Foundations of Animal Behavior*, 1924, 269.
[36] *Ibid.*, 238.

petent neurologist would, and would contend that he has merely presented a schema for clarifying what he means by correlation. But he has nowhere made this admission, and, although schematization may be condensed or simplified, it should never be deliberately misleading. This one is misleading because Herrick has nowhere envisaged this 'tasting right' or 'feeling right' as a problem. On the motor side, one might be led to believe that such a 'simple' reflex activity is one which needs only to be set off by some sort of a trigger. But even limiting ourselves at the moment to external conditions, in the simplest of animals possessing a nervous system such a food-getting activity varies with the pattern of excitation and also with patterns set up within the nervous system.

It would seem unnecessary to dwell at such length upon a single illustration, were it not that it carries the implication that knowledge of such a schema for simpler activities may aid us in understanding the more complicated. Herrick sets up as the primary functions of the nervous system those of integration and correlation, and then gives us this picture to illustrate how correlating mechanisms work.

In short, the primary objection to Herrick's *Neurological Foundations* is that he has not broken away far enough from traditional modes of approach. He has sketched the outline of a new approach to the functions of the nervous system. How much promise such an approach holds we have little opportunity to judge, since it is so often lost in a mass of traditional theories. He has attempted the impossible task of taking account of most of the current theories along with this new approach.

In the volume called *Brains of Rats and Men*,[27] Herrick has dealt more specifically with the cerebral cortex; in fact the subtitle is *A Survey of the Origin and Biological Significance of the Cerebral Cortex*. With such a subtitle, we are not surprised to find that the volume contains a wealth of information about the comparative structure of animal brains. By the reasoning discussed earlier, Herrick has arrived at, and has been much impressed by, the conclusion that the cortex is chiefly concerned with learning and memory, so he has also devoted several chapters to a summary and criticism of Lashley's work up to 1926.

Herrick lays the foundation for his discussion of learning with an account of the conditioned reflex.[28] Conditioning is explained in terms of *drainage*, or the lowering of resistance in central pathways, and we are shown the familiar switchboard arrangements to illustrate what occurs in the nervous system. Lashley had previously criticized such explanations in terms of drainage, or indeed any doctrine attempting to explain in terms of specific neural pathways. In support of his argument he had pointed to instances in which monkeys had learned an activity with one hand while the other was paralyzed. Upon recovery, this formerly paralyzed

[27] C. J. Herrick, *Brains of Rats and Men*, 1926, 1-382.
[28] Herrick, *ibid.*, Chap. III.

hand was used in the activity with as great facility as the other had been. Lashley contends that this shows that pathways not used at all during the original learning may be utilized in later carrying out the activity.

Herrick says in reply:

"This is a gratuitous assumption, for there is an immense complexity of neural architecture in addition to that of the precentral gyrus which was certainly involved in this training and which may have been entirely unaffected by the operations, and drainage or any other process which one wishes to hypothecate may have been involved in the transfer of the training to the other hand. These monkeys had damage done only to a sector of the final common path which may be utilized not only by this habit, but by countless others. The engram of the particular problem-box habit in question is certainly not in the precentral gyrus (which alone was injured); it must lie in some premotor arrangement of associational neurons not at all affected by the operations."[39]

This quotation should serve to show the elusive nature of these processes of 'drainage.' If we can show that they cannot occur in one portion of the nervous system, then, since they exist *a priori,* they must occur somewhere else. Such a defense eliminates the possibility either of proof or of disproof, and requires that we accept them on faith.

It would be profitless to take up Herrick's extended treatment of Lashley's experiments on learning, since it covers only Lashley's work before 1926, and since most of the major criticisms would require revision in the light of Lashley's more recent work.

In the rest of the book, even though we keep in mind that this is a book written for the reader little versed in the details of neurology, we find problems stated and treated in such gross terms as to cover up the dearth of well-supported knowledge. For the more sophisticated reader, there is little more than is to be found in any elementary text in neurology, *plus* some broad statements about 'mind' and 'consciousness.' To be more concrete, Herrick points out (Chapter XII) the fact that the anterior peduncle of the thalamus connects with the frontal cortex, and that the thalamus is supposed to be the "organ par excellence of affective experience." From this he concludes that it must be in the frontal cortex that we have emotional 'drives' knit into patterns of action.[40] Of what use is such piling up of shaky supposition upon shaky supposition?

In the same chapter we are also told that 'phasic' innervation of the motor cells of the cord is mediated by the pyramidal system, and that the extra-pyramidal system probably has to do with 'tonic' innervation. The application of such terms to describe the activities of these systems poses more questions than it solves. No mention at all is made of the tremendous complexity of the latter system, or of the anatomical relationships to the corpus striatum, cerebellum, red nucleus, and other structures. But if Herrick were to lean as heavily upon anatomical evidence for his inferences here as he does elsewhere, the lumping of this whole system under 'tonic' control would scarcely do justice to the anatomical and clinical information available, incomplete and conflicting as it is.

The last chapters of the book, dealing with consciousness and the subconscious, are primarily a plea for dealing with these things as functions of the nervous system. The reader for whom this last portion of the book is apparently designed may find

[39] *Ibid.*, 64.

[40] *Ibid.*, 232.

this a novel mode of treatment of these topics, but the speculative and anecdotal character of the 'psychology' of these sections does not make it clear what sort of a contribution a sound psychology could make with such an approach. In fact, the reader might be led to the notion that the only way of arriving at such a treatment is through neurological studies.

Attempts to arrive at a working conception of the cerebral cortex which will have any significance for the student of psychological functions have only too often been made by those whose knowledge is limited to highly speculative theories of neural function. Herrick, on the other hand, is one of the few neurologists who has tried to take a comprehensive view of the nervous system without burying himself too deeply in anatomy. His use of psychological facts is, however, so broad and so confused that he has glossed over, or quite ignored, many real and pressing problems.

III

In examining the investigations of Lashley and Herrick, our purpose has been to glean from them what we could in the way of anatomical and physiological facts which stand related to the psychological functions of the organism. As psychologists we wish to know not only the way in which the organism operates psychologically, but also the mechanisms which support these functions. It is to the body at large, and to the nervous system in particular, that we must look for these mechanisms.

Lashley's studies illustrate what may be accomplished by limiting a problem to one specific, concrete question, and by treating the results in terms of this question. But the value of Lashley's studies lies not so much in the specific results found as it does in furnishing an example of careful and honest work in this field. As a result of his studies we know something more of the importance of the spatial outspread of the cortex in certain performances of the rat and monkey. We know more of its importance, but little more of its significance. We know, for example, that a certain small region of the rat's cortex is indispensable for behavior depending upon discrimination of visual patterns of a spatial nature. But this tells us nothing of the essential mechanisms involved in the cortex when the organism is carrying on such a performance. Lashley's results also indicate that the same region of the cortex may operate in different ways in different sorts of performance, but the unknown nature of these different ways of operating is not revealed by his method, nor does he attempt to build unfounded theories to account for these ways.

Indirectly Lashley's results give us a basis for considering current neural theories of learning untenable from a neurological point of view. Previously

there had been reason to regard these theories as poorly descriptive of psychological facts, but there had been no contradiction on neurological grounds.

From more general works—such as the two books of Herrick, in which an attempt is made to present a general view of neural function—we find it still more difficult to glean material concerning the essential functional modes of the nervous system. We must use extreme caution in sifting out our material here, because there is a danger that in these general works certain labels will lose their original flavor as descriptive of experimental facts or items of observation, and take on a new and quite unfounded signification. We have pointed out a few instances of this sort in Herrick's writing as examples of the way in which an illicit change of meaning may mar the most forward-looking speculative writings on the nervous system. The term 'plastic' is used in such a way as to imply that it describes cerebral mechanisms, whereas it is really descriptive of a characteristic of behavior. The terms 'phasic' and 'tonic innervation' are supposed to tell us something of the functions of two great motor systems, whereas they are rather terms which describe diverse types of muscular contraction.

We may see from these instances that change in the signification of such terms may lead the unwary reader to quite unwarranted assumptions about neural function as it might be experimentally determined. We may also see that the shifts in application of the terms usually come about either through doubtful assumptions concerning the meaning of experimental results or else through unclear thinking.

Caution against such errors is not of course to be exercised by psychologists alone, but psychologists should be particularly on their guard against these errors since they arise very often when psychological terms are taken over by students of related sciences.

The neurologist does not recognize the gaps in his knowledge of psychological performance, since he has a number of psychological words at his command which he uses uncritically. Thus 'learning' (used loosely to include memory) becomes a faculty of profiting by repetition. The vast complexity of changes which come about when a performance is learned, and the complexity of the factors influencing these changes, are then ignored. In the previous illustration, where certain movements are described as *per se* 'voluntary,' the authors really leave the psychological problems of willed action entirely out of account, yet when they investigate the functions of the precentral cortex they feel wholly justified in speaking of the importance of this area in 'willed' activities.

It is for reasons similar to these that we find difficulty in deriving from

Herrick the exact bodily resources of psychological performance. Among his numerous references to psychological terms, we find very little that deals with concrete psychological problems. If we select a specific problem, such as that of perceiving, we find much material on the manner in which the sense organs are connected to the brain and of the structure of the nuclei with which the receptors are connected. Even this is often oversimplified, as we have pointed out, and it is assumed that such a description tells us how the body perceives, whereas it tells us nothing of the kind.

We have limited our discussion to the points of contact between neurology and psychology, and have criticized neurological discussions only where we find that the neurologist is likely to lead us astray through a lack of understanding of psychological problems, or through a misapplication of psychological terms. So, of course, when the neurologist keeps to neural terms and concepts (leaving psychology out of his reckoning) we have nothing to say. It is only where the language of neural anatomy and physiology is mixed up with a psychological description of performance, or where the behaviorist's stimulus and response, or a learning, thinking, and willing 'mind' are substituted for an accurate treatment of the conditions of psychological performance that we must object.

If the neurologist is to give us answers to questions concerning the nervous system in which we as psychologists are interested, he must be able to understand our quite specific problems and to seek quite specific solutions to them. Such specific queries presuppose a knowledge of the nature of the psychological performances of the organism, and of the conditions which influence them. As long as the neurologist deals in vague, uncritical and antiquated terms, even for such basic performances as perceiving and acting, he will be unable to grasp and subsequently to answer, those questions which will ultimately give us the essential neural operations involved.

THE DIRECTION AND EXTENT OF INTRA-SERIAL ASSOCIATIONS AT RECALL

By JOHN A. McGEOCH, University of Missouri

The problem of the direction and extent of intra-serial associations involves a number of sub-problems each of which is important in its own right and all of which combine to make one of the most theoretically significant issues in the psychology of learning. The more general problem is to discover whether during the learning of a given material, as, for example, a list of adjectives, associations are formed only in the forward direction or in both the forward and backward directions, and, whatever the direction, whether associations are formed between items which are spatially remote from each other in the series. The first part of the problem is that of the direction of associations, while the second is that of remote association. Each may be studied during practice or during recall after practice has ceased. The present experiments have studied the two aspects of the problem at recall.

Granted the existence of remote associations, in particular, it is desirable to know whether their formation represents a basic feature of serial learning or whether it is specific to the *S*'s mode of attack. In all serial learning the non-adjacent items are inevitably remote in space from every item to which they are not adjacent, but it does not follow that they are also temporally remote in the reaction series of the learner. He may bring spatially isolated terms together representatively and react to them as to a sequent pair or, if all of the series is before him at once, he may react directly to isolated items as if they were in sequence. Psychologists have usually assumed that the fundamental question is that of association between temporally remote terms, and there is little doubt that it is the more fundamental, although the other question of the formation of associations between spatially remote terms by any devices whatever has more importance than has usually been thought.

The conventional method of studying the problem with verbal materials has been that of the test list or derived list whereby a series is learned in the usual manner, then rearranged in some way to form a test list and relearned. The rearrangement may involve placing new items at alternate positions on the assumption that, if remote associations have been formed between items 1, 3, 5, etc., this list will be easier to learn; or it may involve the relearning of the series in the reverse

* Accepted for publication October 15, 1934.

direction or in many different forms of combination.[1] This method has the disadvantage that the learning of the derived list involves unmeasured amounts of associative inhibition and that remote associations, if present, may be entirely masked or, if they appear, can be but a residual left after inhibition has been cancelled.[2] It measures, in addition, only the transferred effects of remote associations acquired in the original list and later applied to the learning of the test list. It does not measure either the direct formation of remote associations during the learning of the original list or their individual presence during recall and during the learning of the derived list. Work by this method has frequently found evidence for remote associations in both directions, but has not established whether or not they are temporally remote. What data there are on the question point to an inverse relation between amount of association and degree of remoteness.

There have been, also, two groups of experiments upon the direction and extent of associations during learning itself. One group has been done largely by conditioned response methods and has yielded positive evidence for the formation of associations in both directions, together with some evidence of remoteness.[3] The second group has studied anticipatory responses and has found remote associations in the forward direction.[4]

A fourth general method has been employed by Wohlgemuth to test for remote associations at recall.[5] Series of simple diagrams or of syllables were presented serially

---

[1] The major work on the problem, prior to 1933, has been listed elsewhere and need not be cited again. Cf. J. A. McGeoch, The psychology of human learning: A bibliography, *Psychol. Bull.*, 30, 1933, 1-62, esp. 15-16; and Learning and retention of verbal materials, *ibid.*, 31, 1934, 381-407.

[2] Let a series of 10 verbal items, A to J, be learned and let remote associations over one skipped term be tested by the method of derived lists. Consider the case of the method in which the derived list is made up by inserting a new item between alternate items. In the learning of the derived list old item A must now be connected with a new item, N, and N must be connected with old item C, and so on. A has already been connected with B and B with C. The connection of A with N and of N with C offers opportunity for associative inhibition. If the items inserted are from another list previously learned by *S*, the opportunities for associative inhibition are doubled. If associative inhibition is present in any amount, it will mask remote associations, if any have been formed. If the old items are removed from their original positions, there is further opportunity for associative inhibition. If they remain in their original places, place association may give a spurious appearance of remote association, if strong enough to overbalance associative inhibition. The method of derived lists involves other variables, such as transfer from list to list, which are very difficult to control.

[3] For an introduction to this work see E. S. Robinson, *Association Theory Today*, 1932, 1-142; W. S. Hunter, Basic phenomena in learning, *J. Gen. Psychol.*, 8, 1933, 299-316; and E. R. Guthrie, Association as a function of time interval, *Psychol. Rev.*, 40, 1933, 355-368.

[4] Cf. F. H. Lumley, An investigation of the responses made in learning a multiple choice maze, *Psychol. Monog.*, 42, 1931 (no. 189), 1-61; Anticipation of correct responses as a source of error in the learning of serial responses, *J. Exper. Psychol.*, 15, 1932, 195-205; Anticipation as a factor in serial and maze learning, *J. Exper. Psychol.*, 15, 1932, 331-342; R. Dodge, Anticipatory reaction, *Science*, 78, 1933, 197-203; M. B. Mitchell, Errors in the memorization of numbers, this JOURNAL, 45, 1933, 1-16; and Anticipatory place-skipping tendencies in the memorization of numbers, this JOURNAL, 46, 1934, 80-91.

[5] A. Wohlgemuth, On memory and the direction of associations, *Brit. J. Psychol.*, 5, 1913, 447-465.

and learned to one correct reproduction. Single items were then presented as stimuli and the *S* was to react as soon as an item immediately preceding or succeeding the stimulus or next but one to it had been recalled. When the direction of recall was not specified, there were more forward than backward recalls and the forward dominance was more marked with syllables than with diagrams. This difference decreased when the *S* had to pronounce "ten" aloud every time a syllable appeared in the window. Specification of the direction of recall had a marked positive effect.

The data to be reported here have been gathered by a modification of Wohlgemuth's method. In the present experiments the *S*s have been left free to give whatever association came first. It is possible, thus, to test for associations of all degrees of remoteness, whereas by Wohlgemuth's method the search was limited to adjacency and one degree of remoteness. An effort has been made to obtain a picture of the direction, extent, and certain conditions of intra-serial associations as they appear under the association-method, and of the extent to which intra-serial associations are stronger than extra-serial.

The association-method has certain advantages. It is free from many of the masking effects of the method of derived lists and it affords a more detailed and immediate picture of the associative connections within the list than does any one of the other methods. It involves, of course, a change in stimulating conditions. The method introduces a change in instructions and in set, together with a change in the mode of presentation of the stimulus, the performance of the E, and in other stimulational details. These changes may operate to decrease retention and they may possibly inhibit remote associations; there is no reason to suppose that they act to increase spuriously the amount of remote association. The *S*s were, moreover, practiced upon two lists, with the change in stimulating context, before the beginning of the experiment proper and had, therefore, become partially adapted to it. The results may have been affected to some degree by the change from learning to testing, but the influence could only have been to mask some remote associations; the positive results could not have been produced by the change.

Experiment I

The data upon remote association, which are to be described under the heading Experiment I, were gathered during the course of an experiment upon frequency of repetition *vs.* amount of interpolated material as determiners of degree of retroactive inhibition.

Source of the data. The learning materials were lists of 10 two-syllable adjectives, plus an initial adjective which served as a cue to instigate the first word of the main list but which was not itself learned as a response. The lists were exposed at the rate of 1.8 sec. per word on an electrically driven drum and were learned by the anticipation method. In each of the 5 conditions of the experiment an original list was given 5 learning presentations followed by a 10-min. interval variously filled. At the end of this interval the list was relearned or, since in most cases learning had not been completed to even a low criterion during the 5 learning

presentations, learning was continued to a criterion of 3 perfect successive repetitions.

The conditions of the experiment consist of one rest and 4 work conditions. The only difference between the 5 conditions lies in the filling of the 10-min. interval. Under the rest condition *S* read jokes and marked the 3 best on each page. Under each of the work conditions *S* received 12 presentations of interpolated material of the same kind as that constituting the original lists, but the number of lists and the frequency of repetition of each list varied thus: Work-1, 12 repetitions of 1 list; Work-2, 6 repetitions each of 2 lists; Work-3, 4 repetitions each of 3 lists; and Work-4, 3 repetitions each of 4 lists. Where more than one interpolated list was learned *S* passed from one to the other with no more break than from one trial to the next of a single list.

The 25 *S*s went through each of the 5 conditions in a counterbalanced practice order at the rate of one condition per day. The *S*s were college students without prior practice at the anticipation method. Before the beginning of the experiment each was given a practice sitting at which he learned and relearned 2 lists. The experiment was conducted by an *E* with 3 semesters of practice in the conduct of experiments upon retroactive inhibition.[6]

Method of obtaining the data. As soon as each *S* had reached the relearning criterion of 3 perfect successive trials he was given the following instructions:

"Keep the list of words you have just learned in mind. I shall call some words at random from this list and I want you to close your eyes and give me the first word that comes to your mind in response to each word I call out. The word you first think of will probably be a word from the list, but whether it is from the list or not give me the first word that comes to your mind."

E said "Ready" before each stimulus-word. The association-times were taken by a stopwatch graduated in hundredths of a second. Associations were asked for after the relearning of the practice lists so that the *S*s were accustomed to the method before the data reported were collected.

In this experiment, the first and tenth words were not given as stimulus-words. The words at the eight other positions were presented in 5 different orders, one for each position in the practice order of the lists, so that the frequencies of the orders are the same for all conditions of the experiment. The records for one of the *S*s upon one of the lists are not usable. The total number of records, therefore, at each serial position is 124 and the total number for all lists and positions is 992.

Scoring methods. The response to every stimulus-word can readily be given a value in terms of its serial distance from the stimulus. The unit employed is the number of positions skipped. If the response-word follows the stimulus-word in the list, it is classified as a case of zero forward association; if the response is from the second forward position, it is classified as representing one degree of remoteness in the forward direction, and so on. The same procedure has been used with associations in the backward direction.

Since only the words from positions 2 to 9, inclusive, have been used as stimuli, it is possible for every word to elicit an associate from either the immediately following or the immediately preceding position. If we assume that each list is learned

---

[6] This experiment and the following one were assisted by a Grant-in-Aid from the National Research Council. The first three experiments were conducted by Mr. Willis H. McCann.

as a linear series which begins with the cue word and ends with the word in position 10, the greatest possible degree of remoteness in the forward direction is 7 positions, *i.e.* from position 2 to position 10. The corresponding degree of remoteness in the backward direction, however, is 8 since the initial or cue word may be given as response to the stimulus word in position 9.

TABLE I

NUMBERS OF FORWARD ASSOCIATIONS
(On the assumption that a list is a linear series)

Degree of remoteness	Serial position of stimulus-word								Total
	2	3	4	5	6	7	8	9	
0	23	22	34	11	27	25	34	35	211
1	7	8	7	10	16	8	12		68
2	4	7	5	9	12	7			44
3	5	7	11	11	5				39
4	10	3	4	15					32
5	3	17	18						38
6	8	10							18
7	12								12
Total	72	74	79	56	60	40	46	35	462

The assumption of linearity is in the present instance a very reasonable one. The *Ss* were told the number of trials to be given during the original learning and this instruction would lead them to treat each trial as a separate repetition of a linear series. Each list had, moreover, an initial cue-word which was not to be anticipated on each trial when the last word of the list appeared. When the last word had been in the window of the screen for 1.8 sec., the drum was reset with the cue-word in the window to indicate the beginning of another trial. This break between trials would emphasize further the linearity of the series.

The assumption that each list is for *S* a linear series is not, however, the only one that can be made. When the successive presentations follow each other with only a 5-sec. interval between, it may be that the list is learned as a circular series in which the last term becomes the stimulus to recall of the cue word at the beginning even though the cue is not called for as a response. The most defensible method of scoring the results, on the assumption of serial circularity, is to classify an associate as forward or backward on the basis of minimal serial distance. Thus, if a stimulus-word from position 8 elicits a response-word from position 2, this is called an association 4 positions (terms skipped) remote in the forward direction instead of 5 in the backward. By this method of scoring there are only 5 possible degrees of remoteness in either direction, viz., 0, 1, 2, 3 and 4. The data have been treated in terms of each assumption.

Results. The common assumption is that lists of items learned by the anticipation method are learned as linear series and the data first to be presented have been classified on this assumption. The number of associations of each degree of remoteness in the forward direction is shown in Table I.[7] When the lists are assumed to be linear, the extent of the

[7] In this table and in the following ones "serial position of stimulus-word" means the position occupied by that word in the list during learning and without regard to the

possible remoteness of the association given decreases linearly from the second to the ninth position of the stimulus word, and this must be remembered during the interpretation of the results.

It can be seen at once that by far the greatest number of associates in any single classification comes from the immediately following position (211 cases). This number comprises 45.67% of all the associates given in the forward direction. The noteworthy thing, however, is not that so large a percentage of the associates comes from the adjacent forward positions but that the percentage is no larger. The lists were learned so that each adjective elicited the next succeeding adjective, yet when *S* is given an opportunity to associate freely, over half (54.33%) of his forward associations are one position or more than one removed from the stimulus-word. The fact of remote forward association under these conditions of learning and of testing is, thus, established.

We shall inquire, next, into the relation between frequency of remote forward association and degree of remoteness. This relation cannot be discovered by an examination of the totals at the right side of Table I. The inverse relation between these totals and degree of remoteness is a function of the fact that the upper limit of the possible remoteness of association decreases linearly as the position of the stimulus-word moves from the beginning to the end of the series. One must examine, instead, the numbers of associations at each serial position to see if they show any regular variation with remoteness. The second serial position has the widest range of possible remote associations while each succeeding position loses one such possibility. At no position is there any consistent relation between degree of remoteness and number of associations, and evidence for the obvious inference from Ebbinghaus' results, viz., that the relation should be inverse, is entirely lacking.

The only variation which appears with any consistency is that in the direction of an increased number of associations at the greatest or next to the greatest possible degree of remoteness from each of the first 4 positions of the stimulus-word. This relatively large number of associations between each of the first 4 terms and the final term in the series may imply a kind of pyramiding of the list toward the last item, in terms of remote associations, or it may be interpreted as a case of less remote backward association on the assumption of the circularity of the series. It is neither sufficiently large nor sufficiently regular to merit detailed examination.

Related to this is the question of the variation in remote association with the serial position of the stimulus-word. Inspection of the table shows a tendency for

---

order in which it was given as a stimulus during the test. The words were presented, in testing for remote association, in 5 different orders, each different from the order during learning.

the later positions to yield a disproportionately large number of associations. This is especially true when only immediately successive associations are considered, but it is also true with other degrees of remoteness. The totals at each serial position do not decrease in size at the rate at which the possibilities of varying degrees of remoteness decrease; they even increase from serial position 2 to serial position 4, and the decrease thereafter is irregular. This result may be in part a more or less necessary consequence of the decreasing opportunities for remote associations at the later positions and the resulting formation of a relatively larger number between the remaining terms. It is probably not a result of the decreasing opportunities alone.[8]

TABLE II

MEAN ASSOCIATION-TIMES (IN SEC.) OF THE FORWARD ASSOCIATIONS LISTED IN TABLE I

Degree of remoteness	Serial position of stimulus word								Mean, all positions
	2	3	4	5	6	7	8	9	
0	1.65	1.39	1.85	2.71	2.00	1.63	2.18	1.21	1.76±.12
1	1.44	1.24	1.89	1.59	1.41	1.29	1.47		1.46±.09
2	1.82	1.24	1.92	1.35	1.33	1.48			1.45±.11
3	1.37	1.08	.93	1.17	1.06				1.09±.09
4	1.25	2.62	1.43	1.45					1.55±.19
5	1.23	1.18	1.28						1.23±.15
6	1.19	1.42							1.31±.14
7	1.21								1.20±.14

The association-times of the responses from different degrees of remoteness provide one measure of relative associative strength since it is reasonable to assume that association-time and associative strength are inversely correlated. The mean times of the forward associations of Table I are presented in Table II. The sigmas in the column at the right are sigmas of the means. On the assumption that the word most strongly associated with any given word is the one which follows it and which has been its response-word throughout the learning repetitions one would expect the association-times at zero remoteness to be shorter than those at any other degree. The further assumption, which has also been established for some time as the conventional one, that the strength of remote associations varies inversely as degree of remoteness requires that the association-times should increase from zero remoteness to whatever is the greatest degree of remoteness possible at any given serial position.

The implications of neither assumption are justified by the data. The mean times when the response-words are those which immediately follow the stimulus-words in the list are in a majority of the cases longer than

[8] It is possible, by making certain assumptions, to compare the expected and obtained frequencies at each serial position, but in view of the relatively small frequencies at the higher degrees of remoteness little would be gained by such a comparison beyond what can be seen from an inspection of the table.

any of the others.[9] The mean time for all serial positions at zero remoteness is larger than the mean at any other degree. The differences between the mean at zero remoteness and at other degrees are statistically reliable in only a few cases, as in those of degrees 3 and 7, but the uniformity of the direction of the difference implies its trustworthiness. From zero to the greatest remoteness at each position no wholly consistent variation in time takes place and there is no hint of an increase in time with increasing remoteness. There is, instead, an irregular decrease. The rank-difference correlation between degree of remoteness and mean association-times at the second serial position, which has the widest range of remoteness, is —0.90 and while the inverse correlation is not as high at the other serial positions the relationship only once becomes positive.[10]

The association-times at zero remoteness have a tendency to be larger at the intermediate serial positions than at the initial and final ones. This suggests the conventional serial-position curve obtained when rapidity or degree of learning is plotted against position. The same result appears at one-term remoteness. The progressive shortening of the range of positions at which entries are possible forbids any conclusion regarding the relation at the higher degrees of remoteness.

It is interesting that the mean times for most degrees of remoteness and at most serial positions are shorter than the 1.8 sec. within which correct anticipation was required during learning. The mean times for all serial positions combined are uniformly less than 1.8 sec. The *S*s were able, on the average, as the times show, to recall under the changed conditions within the time limits set during the learning.[11] This fact is strong evidence that the change in conditions from learning to test has not greatly affected the *S*s.

We have, thus far, been treating all forward associations as belonging to a linear series. If, on the other hand, we assume that the list is circular and count the direction of associations by the shortest route, we obtain a slightly different result. This method of counting yields no regular variations with the serial position of the stimulus-word and the results at each

[9] In only 5 of the 28 entries recording remoteness of one intervening term or of more than one are the mean times higher than those for zero remoteness and in 3 of the 5 the difference is insignificantly small.

[10] The coefficients at the next three positions are 0.13, —0.60 and —0.70, respectively. N is, of course, too small to permit these coefficients to have statistical importance.

[11] There are four exceptions to this in the means for zero remoteness at different serial positions and there are occasional exceptions at other degrees. The mean times at all serial positions combined are, however, the more stable and the more representative.

position may be omitted. The total number of associations and the mean time thereof at each of the 4 degrees of remoteness which are possible according to this method are shown in Table III. The sigmas are sigmas of the means. In order to conserve space the corresponding data upon backward association are given in the same table. Reference will be made to them later. Associations of zero remoteness are the same by either method of counting and are omitted.

TABLE III

NUMBER OF FORWARD AND OF BACKWARD ASSOCIATIONS
(On the assumption that a list is a circular series)

Degree of remoteness	No. forward associations	Mean time (Sec.)	No. backward associations	Mean time (Sec.)
1	71	1.47± .09	52	1.46± .09
2	57	1.45± .11	62	1.47± .10
3	59	1.14± .08	49	1.31± .10
4	48	1.43± .13	64	1.25± .06
Total	235	1.37	227	1.37

The largest number of remote forward associations occurs across but one intermediate term and there appears a tendency to a decrease in the number as more terms are skipped. The association-times do not vary regularly or importantly.[12] To which way of counting one should give the greater weight is, in the absence of a satisfactory objective criterion, a matter of individual judgment. In an experiment to be described later, in which all of the items were used as stimulus-words, a few associations from the last word of a list to the cue word have been obtained and it is reasonable to interpret these as cases of adjacent or zero degree forward association rather than as cases of 9 degrees of remoteness in the backward direction. This assumes circularity of serial learning. For the sake of completeness it is desirable to treat the data on the basis of both assumptions, although the data obtained on the assumption of linearity should carry the greater weight. There is, however, scant need to argue the point since the results of the two methods are in complete agreement except upon the single point of the relation between frequency of association and degree of remoteness and there the disagreement is not great.

The 462 forward associations of Table I make up only 46.57% of the total number of associations elicited. Of the remainder, 359 or 36.19% are in the backward direction. Their distribution is shown in Table IV. Eight degrees of remoteness are listed because in 3 cases the stimulus-word in the ninth serial position elicited as response the cue word of the list.

[12] The difference between the means at 1 and 3 degrees of remoteness is 0.33 ± .12. It is doubtful whether any significance should be attached to this difference, in spite of its approach to statistical importance.

The largest number of backward associations at any single degree of remoteness is at zero or adjacency. At this degree are 40.9% of all the backward associations given. The striking thing, however, is that 59.1% of such associations are from more remote positions in the list. The fact of backward association under these conditions is indubitable.

TABLE IV

NUMBER OF BACKWARD ASSOCIATIONS
(On the assumption that a list is a linear series)

Degree of remoteness	Serial position of stimulus-word								Total
	2	3	4	5	6	7	8	9	
0	19	21	11	27	19	23	12	15	147
1	7	8	4	5	10	4	7	7	52
2		3	7	5	4	12	9	10	50
3			3	5	4	7	6	6	31
4				4	9	3	7	4	27
5					2	4	5	5	16
6						7	5	8	20
7							5	8	13
8								3	3
Total	26	32	25	46	48	60	56	66	359

The relation between frequency of backward associations and degree of remoteness can be determined by inspection of the columns at each position. There is an irregular tendency for the number of associations to decrease as remoteness increases. The rank-difference correlation at the ninth position is —0.52 and a tendency toward an inverse relation can be seen to exist in the other columns. The decreases from one degree of remoteness upward are, however, relatively small and do not represent at most positions the amounts which one would be led by past work upon the problem to expect.

The frequency of associations of zero remoteness is somewhat greater at the intermediate positions, especially at 5 and 7, than at the end positions. The differences are small but the fact that the intermediate positions yield even as many backward associations as the better learned initial and final positions is of importance. No other noteworthy variations with position are discernible.

Table V gives the mean association-times of the associations listed in Table IV. The mean times at zero remoteness are not consistently different from those at other degrees and there exists no consistent relation between degree of remoteness and association-time, although there is a tendency for the times to grow shorter with greater remoteness at some positions. The differences between the means in the last column are, however, uniformly unreliable. The only legitimate conclusion is that no important differences

have been found between the association-times at different degrees of remoteness.

When the number of backward associations at each degree of remoteness is counted on the assumption of the circularity of the list (Table III), no indication of a consistent relation between frequency and remoteness is

TABLE V

MEAN ASSOCIATION-TIMES (IN SEC.) OF THE ASSOCIATIONS LISTED IN TABLE IV

Degree of remoteness	Serial position of stimulus-word								Mean
	2	3	4	5	6	7	8	9	
0	1.08	1.36	1.55	1.21	1.35	1.31	1.46	1.21	1.29± .08
1	1.37	1.28	2.56	1.24	1.24	1.28	1.55	1.63	1.46± .09
2		.99	.95	1.18	1.64	1.33	2.33	1.66	1.51± .12
3			.83	.83	1.20	2.01	1.15	1.31	1.30± .11
4				1.13	1.09	1.13	1.30	1.18	1.17± .10
5					1.15	1.13	1.28	1.11	1.17± .14
6						.84	1.20	1.48	1.18± .13
7							.90	1.74	1.43± .26
8								1.59	1.59± .57

to be found. Mean time, however, decreases by statistically unreliable amounts with increasing degree of remoteness.

The sum of the forward and backward associations is 821. The remaining 171 associations, out of the total number of 992, were not in the lists learned. The *S*s had, that is, under these conditions of recall, a stronger association between 171 of the words and other words than those which had just been learned, in spite of the fact that the list had been brought as a whole to a criterion of 3 perfect trials in succession. These associative responses, which comprise 17.24% of the total, had a mean association time of 2.14 $\pm$ .14 sec. They required, thus, a considerably longer time than either forward or backward associations from within the list.

Comparison of forward and backward association. The data have shown conclusively that under the conditions of this experiment both forward and backward associations of more than zero remoteness are formed in considerable numbers. It remains to compare the results upon the two different directions. The data for such a comparison are already available in Tables I to V.

We shall compare first the relative frequencies of forward and backward associations and shall treat zero remoteness separately from the other and positive degrees of remoteness. The numbers of forward and backward associations at the zero degree are 211 and 147 and the ratio of backward to forward is 1:1.43. When the frequencies at the other degrees are summated on the assumption of linearity of list, the backward-forward ratio is 1:1.18, and when treated on the assumption of circularity of list,

the ratio is 1:1.03. There is, thus, a consistent difference in favor of recall in the forward direction, and this difference is smaller when the results at the positive degrees of remoteness are placed in comparison than when those at zero remoteness are compared. The magnitude of the difference is a function of the method of counting direction.

The two directions of recall have relative association-times which are not to be expected from work by other methods. At zero remoteness the ratio of backward association-times to forward is 1:1.36. The corresponding ratio for the sum of the times at all of the positive degrees of remoteness is 1:1 regardless of the assumption used in counting the direction.

The frequency of the forward associations does not vary consistently with degree of remoteness, while that of the backward associations shows a tendency to decrease as remoteness increases. The association times in the forward direction decrease irregularly as remoteness increases, and those in the backward direction have a similar tendency although to a very much reduced degree. Forward associations are slightly more numerous between the first 4 positions and the greatest or next to the greatest possible degree of remoteness therefrom. Backward associations have no noteworthy relation to serial position.[13]

Experiment II

The data to be described here were collected by the same method as the one employed in Experiment I but in connection with another experiment upon retroactive inhibition.

Source of data. The experiment involved 8 conditions in each of which the *Ss* learned lists of 16 adjectives, plus an initial cue word, to 8 repetitions, spent the succeeding 20 min. variously, and then relearned the original lists to a criterion of 3 perfect trials in succession. The various interpolations have had no noticeable influence upon the remote association results and may be disregarded. The collection of data upon remote association from 9 *Ss* under the conditions of this experiment represents a preliminary attempt to discover the conditions which affect the phenomenon. While there are other differences than length of list alone between the two experiments, data from lists of 16 items are worth reporting briefly.

Results. In most cases each of the 16 items was presented as a stimulus-word and a total of 1035 associative responses was obtained. The small number at each serial position and degree of remoteness, however, makes it useless to tabulate the results in detail and only a summary will be given.

[13] The data have been examined for differences in remote association between the 5 conditions of the retroaction experiment in which these data have been gathered. None has been found, and the data from lists learned and relearned under all 5 have been combined.

This is especially true in view of the fact that the general outlines of the results are the same as those for lists of 10 adjectives.

The greatest number of forward associations in any single classification is again that at zero remoteness where the number is 170 or 44.0% of all forward associations. The corresponding percentage in Experiment I was 45.7%. The remaining 216 forward associations, counted on the assumption of linearity, are distributed over the other degrees of remoteness from 1 to 14 in a fashion which permits no other conclusion than that frequency is not a function of the number of intermediate terms. Regular variations with serial position are absent. The association-times at zero remoteness have a mean of 1.47 ± .07 sec. as compared with a corresponding mean of 1.76 sec. in Experiment I and those for all other degrees of remoteness average 1.65 ± .07 sec. as compared with 1.36 sec. in Experiment I. The results counted on the basis of circularity of list are the same in tendency.

There are 304 associations in the backward direction and of these 107 or 35.2%, as opposed to 40.9% in the preceding experiment, are the first preceding word. The remaining 197 associations are distributed irregularly over the 15 degrees of remoteness in a way to make inevitable the conclusion that frequency is not highly correlated with degree of remoteness. No consistent variations with serial position are to be observed. The backward association times at zero remoteness average 1.37 ± .06 sec. as compared with the 1.47 sec. for corresponding forward associations and with the 1.29 sec. for the zero backward associations of the preceding experiment. The mean time for those at all other degrees of remoteness is 1.61 ± .10 sec. which is larger than the mean, 1.36 sec., for Experiment I, but still slightly smaller than the corresponding mean, 1.65 sec., for the forward associations in the present experiment. The assumption of circularity yields results of similar tendency.

The results of the two experiments have, thus far, shown an entire agreement in tendency and in some cases an approach to identity of percentages. The largest difference between the two experiments lies in the percentage of associations which comes from outside the list of which the stimulus-word is a member. The number in Experiment II is 345 or 33.3% as compared to 17.24% in Experiment I. The mean association-time is 1.86 ± .07 sec.

Experiment III

The preceding experiments have had in common an initial partial learning, an interval, and the continuation of learning to 3 perfect repetitions. In all but the rest conditions varying kinds and amounts of interpolated learning have taken place during the interval. The occurrence of the rest conditions in a counterbalanced practice order makes impossible the determinate isolation of the influence of the interpolations upon remote association, although in so far as can be told, this influence, if present at all, is not important. The question, however, whether the results obtained under these conditions may be generalized to situations lacking the interval and the interpolation is necessarily present. Since it is peculiarly desirable to

know whether the foregoing results are specific to the conditions of the retroactive inhibition experiment, the following experiment was performed.

Source of the data. Twenty-four college students, all of them practiced in the use of the anticipation method, learned by continuous repetition a list of 10 adjectives (plus a cue word) to a criterion of 3 perfect trials in succession. In order to prevent the results from being specific to a particular list 4 different lists, chosen from those used in Experiment I, were employed. Each one of the 4 was learned by 6 *Ss*. The conditions of this experiment are the same as those of Experiment I except for the features specific to the retroaction conditions. The same tests for remote associations were made as soon as the criterion had been reached. Each of the 10 items was presented as a stimulus-word.[14]

Results. The total number of possible associative responses is 240, but when these are distributed over each position and degree of remoteness the numbers become too small to merit attention. Only totals will, therefore, be considered. The largest number of forward associations, on the assumption of linearity of list, is from the stimulus-word to the following word in the list. The number is 36 or 48.7% of all forward associations as compared with the 45.7% of Experiment I. The 38 remote forward associations, comprising 51.3% of those in the forward direction, are scattered over the 8 degrees of remoteness with no tendency toward regular variation. At zero remoteness the association-times average 1.75 $\pm$.24 sec., and at all other degrees combined they average 1.64 $\pm$ .23 sec. The corresponding figures from Experiment I are 1.76 and 1.36 sec. In the present experiment the direction of the difference is the same but its amount is less.

Twenty-one, or 44.7%, of the backward associations occur at zero remoteness with a mean association-time of 1.69 $\pm$ .31 sec. The remaining backward associations, or 26%, which are distributed without uniform tendency over the 9 possible degrees of remoteness, have a mean time of 1.84 $\pm$ .26 sec. These times are longer than those of the first experiment, but their tendency is the same. The results computed on the assumption of circularity of list are the same in tendency. The largest difference between the two experiments lies in the number of responses from outside the list. The number in the present experiment is 117 or 49.1% of all responses,[15] and the mean time is 1.91 $\pm$ .19 sec.

[14] The first and last words of these lists were also used as stimuli, whereas only the intermediate 8 were employed in the first experiment. There are, thus, 8 degrees of remoteness in the forward direction and 9 in the backward. There is one more degree of remoteness in the backward than in the forward direction because the cue-word may be and occasionally is given as an associative response to the item in the last serial position.

[15] The total number of responses obtained was 238. There were two failures to respond within a reasonable time.

It may be inferred that the remote associations found in the two preceding experiments are not specific to the conditions of the retroaction experiment. The differences between the percentages of remote association in this and in the other experiments are probably a function of differences in degree of learning and in other variables correlated with the counterbalancing of several conditions in the others and perhaps to a slight degree of other differences in method. The fact of the appearance of large percentages of remote association in both directions in a test immediately after the learning of a single list demonstrates, however, that the appearance of remote forward and backward associations is not specific to the conditions of the two first experiments.

Experiment IV

It is possible, at least, that some of the remote associations found in these experiments may be a result of previously formed connections between the words of the list. The words were common adjectives and many of them may have appeared together in the past experience of the *S*s. The instructions would serve to limit the range of associates to the list learned, for the most part, but what item from the list was elicited by any given word may have been a function of prior associations. If this were true, the data could not be interpreted as evidence for remote association. The factor of prior connection could scarcely account for all of the present results, but it might be important. It can be controlled by using as learning materials items which are almost certain not to have appeared together in the past reaction series of the *S*s. Data upon remote associations within lists of such items will be important, also, over and above their significance as a control upon the factor of prior connection.

Procedure. The learning materials were lists of 10 consonant syllables (plus a cue syllable) of the kind devised and used by Witmer.[10] Each item consists of three consonants in an order which does not form any well known set of initials or have any other known prior associative value. On each of days 1 and 2, *S* learned a list to a criterion of 3 perfect trials in succession and was tested immediately for remote association. After 10 min. he relearned the list. On days 3 and 4, *S* went through either a rest or a work condition of a retroactive inhibition experiment under the same conditions as those of Experiment I, and was tested for remote association immediately after the completion of the relearning. The data from the first two lists have been obtained immediately after learning to a criterion. The learning of these two lists provide practice for *S* and he comes to the conditions

[10] Louise Witmer, The standardization of three-place consonant syllables for use in experimental investigations of memory, *Ped. Sem.*, 46, 1935, 337-359. Experiment IV was conducted by Miss Witmer.

of the retroaction experiment with practice at the same number of lists as in Experiment I. He then goes through conditions similar to those of Experiment I, and is tested in the same way as in that experiment. The data thus provide a check on the influence of the test immediately after learning versus that of the test following relearning after an earlier partial learning. The instructions to the *S*s were the same as those of Experiment I with only the changes made necessary by the changed material.

The four conditions were gone through by 15 *S*s previously untrained in learning by the method used. Each of the 10 items in each list was given as a stimulus-word. The total number of possible associative responses is, therefore, 600. The responses for one *S* on the second practice day are not usable. The total number available is, thus, 590.

TABLE VI

NUMBERS OF FORWARD AND BACKWARD ASSOCIATIONS, WITH THEIR MEAN ASSOCIATION-TIMES

(Witmer consonant syllables)

Degree of remoteness	Direction of association			
	Forward		Backward	
	No.	Mean time (sec.)	No.	Mean time (sec.)
0	148	1.94	50	2.12
1	31	2.17	12	1.47
2	19	2.10	22	1.78
3	12	1.80	13	1.99
4	12	1.73	9	1.89
5	17	2.00	10	2.31
6	9	2.29	5	2.27
7	11	2.65	6	2.99
8	6	2.07	5	2.04

Results. The data for each list in the order in which it was learned have been treated separately in order to discover whether there are any progressive changes in remote association which may be ascribed to practice and whether there is any important variation from day to day of the experiment. The numbers of forward associations on the four days in order are 63, 64, 66, and 72, respectively, and the numbers of backward associations are 35, 32, 31, and 34, respectively. The division into zero order and remote association shows no more tendency toward regular variation than do the totals and the data for the four lists will be combined. There is no regular variation with serial position and only the data for total lists will be presented.

The numbers of associations at each degree of remoteness, together with the mean association-times are given in Table VI. The scoring is on the assumption of linearity of list.

The greatest number of associations comes, as in the other experiments, at the

immediately following position (zero remoteness). At this degree of remoteness come 55.85% of all forward associations. This percentage is roughly 10% higher than the corresponding ones of the preceding experiments. A considerable number of associations appear, however, at each positive degree of remoteness from 1 to 8, and the existence of 44.15% of remote associations in material of this sort is very significant. Since the opportunity for associations of the greater degrees of remoteness decreases from the first to the last item in the list, no adequate statement can be made concerning degree of remoteness and frequency of association. Frequency does not, however, decline with increasing remoteness as rapidly as preceding work on the problem would lead one to expect it to do.

The association-times, both at zero remoteness and at the positive degrees, are longer than those for the adjectives. They are, however, with but one exception, less than a half second longer than the 1.8 sec. exposure-time within which association was required during learning. They have a tendency to be larger at the greater degrees of remoteness but there is no regular variation. They average $2.10 \pm .14$ sec. at the positive degrees of remoteness, as compared with $1.94 \pm .09$ sec. at zero remoteness.

The 265 forward associations make up 44.9% of the total number given. There are, also, 132 backward associations, or 22.4% of the total number. Of these 132 backward associations 50, or 37.88% are at zero remoteness and 82, or 62.12%, are at some positive degree. These percentages are not widely different from those obtained with adjectives. Frequency declines with increasing degree of remoteness, but this is, in part at least, a function of the decreasing opportunity for remoteness of higher degrees as one moves from the end to the beginning of the list.

The association-times of the backward associations are longer than those for adjectives, but they are of the same order of size as are the corresponding times for the forward associations in the present experiment. At zero remoteness the times are 1.94 sec. in the forward direction and $2.12 \pm .12$ sec. in the backward. For all positive degrees of remoteness combined they are 2.10 sec. in the forward direction and $1.98 \pm .13$ sec. in the backward.

When the data are scored on the assumption that the list is a circular series, the results correspond in implication with those obtained on the assumption of linearity. There is no doubt that both backward and forward associations of all possible degrees of remoteness are formed between the consonant syllables, regardless of the assumption used in scoring.

At the tests for remote association 397 items were recalled correctly. In 193 cases the data are, for one reason or another, incorrect. Some combinations of consonants not in any list learned were given; some were from former lists; some reversed the order of the letters of an item or gave them in some other incorrect order; and in a few cases no response at all was given. These cases are not immediately important for the problem and will not be analyzed further. It is useless to give their association-times.

The data of this experiment agree in tendency with those of the preceding experiments. When the difference in material is considered, even the differences in percentages of forward and backward, zero and remote

associations are relatively small. The results corroborate those obtained with adjectives and imply that the results from the adjectives are not a function of previously established connections between the items of the lists.

A further possibility has been considered in connection with these data. The *S*s might give their associative responses, after the first, to the responses which they have just given. The response to the first stimulus-word item might arouse, that is, another item in the list and this item might then be given in response to the next stimulus-word. This could give a spurious picture of remoteness of association. As a check upon this possibility, the data have been scored in terms of the relation of each response to the preceding one. No evidence of association from response to response has been found. If the *S* has been making such associations, they are as remote as those found when scoring is in terms of association from stimulus-word to response.

Discussion

The experimental results have been presented and we may now turn to an examination of their meaning with particular reference at first to the results of Experiment I. In lists which have been exposed and learned to a criterion of 3 perfect successive trials in a uniformly forward direction so that each word in the list serves as a stimulus to the recall of the next following word, and when immediately after the attainment of the criterion each word is presented with the instruction to speak the first word which it arouses, it might be expected that, in spite of the fact that the words are now presented in a new order, the great majority of the responses would be the next following word which has been throughout the learning of the list the response to this stimulus. This is not the case. In Experiment I slightly less than half of the associations given are in the forward direction at all and of these less than half are from the next following position. Somewhat more than half of the forward associations are, thus, remote associations in the conventional sense of associations which are separated by one or more serial positions from the stimulus-word.

These remote associations do not, moreover, show any certain indication of a decreasing frequency with increasing degree of remoteness, and they are given with a rapidity greater than that of words of zero remoteness, a rapidity which gives no hint of decreasing as remoteness increases. These particular results do not fulfil the expectation from the results obtained

by other methods. They cannot, however, be accepted as final on account of the relatively small number of cases at the higher degrees of remoteness. The agreement among the four experiments and the failure of time to increase with increasing remoteness implies that this aspect of the problem merits further study by this method.

The conditions of the experiment are specifically arranged to produce learning in the forward direction, yet when the *Ss* are asked to give the first word recalled by the members of the list, 36.2% of their response-words lie in the backward direction from the stimulus-word. These back-

TABLE VII

SUMMARY OF THE RESULTS OF EXPERIMENTS I, II, III, AND IV
(Time in sec.)

	Experiment I		Experiment II		Experiment III		Experiment IV	
	%	Mean time	%	Mean time	%	Mean time	%	Mean time
Forward Direction:								
Total	46.6	1.54	37.3	1.57	31.1	1.69	44.9	2.01
Zero remoteness	45.7	1.76	44.0	1.47	48.7	1.75	55.9	1.94
All other degrees	54.3	1.36	56.0	1.65	51.3	1.64	44.1	2.10
Backward Direction:								
Total	36.2	1.33	29.4	1.52	19.8	1.77	22.4	2.03
Zero remoteness	40.9	1.29	35.2	1.37	44.7	1.69	37.9	2.12
All other degrees	59.1	1.36	64.8	1.61	55.3	1.84	62.1	1.98
From outside list	17.2	2.14	33.3	1.86	49.1	1.91	32.7	—

ward associations are not between the stimulus and the first preceding word only. On the contrary 59.1% of them are from more remote positions and these remote associations show only a slight and uneven tendency to decrease with increasing degree of remoteness. The mean association-time of these remote backward associations is very slightly larger than that of the zero degree backward and is exactly equal to that of the remote forward associations. The mean time of the zero degree backward associations is considerably smaller than that of the corresponding forward associations. These facts present a picture somewhat different from the one implicit in the current assumed dominance of the forward direction. Recalls in the backward direction, far from being infrequent, occur within 10% as often as recalls in the forward direction, with a greater percentage of frequency at the positive degrees of remoteness and with equal or greater rapidity.

These results have been obtained from lists which were partially learned and then continued to the criterion after an interval. In order to discover whether the results are specific to these particular conditions Experiment

III, in which the list was learned to the criterion by continuous repetition, was performed. To facilitate a comparison of these and of the other experiments a summary is given in Table VII. A few percentages not previously presented have been added to make comparisons uniform. The data given are computed on the assumption of linearity of list. Computation on the assumption of circularity does not alter the conclusions.

The results of Experiments I and III are sufficiently close together in general tendency to permit the statement that the conclusions of Experiment I are not specific to the conditions of the retroactive inhibition experiments. There are differences between the two in percentages and in mean times, but the agreements are more striking than the differences and the uniformity of tendency is the important fact in the present context. The data of Experiment II are, likewise, sufficiently similar in implication to those of Experiment I to indicate that the results are not to any great extent a function of that particular length of list.

The fact that Experiment IV, in which consonants were used as the learning material, yields results of the same general tendency strongly implies that the data of the three experiments with adjectives as learning material are not a function of the specific material or of prior association between the words.

The differences between the mean association-times of the four experiments and of the different classes of associations are in very few cases statistically reliable and it would be fruitless to compare them statistically in detail. The sigmas previously given in Tables II and V and in the description of the results of the other experiments permit the major comparisons to be made if desired. The major conclusions based on association-times do not require detailed comparisons. They are (1) the association-times for forward and backward association do not differ greatly or consistently from each other; (2) both forward and backward associations at the positive degrees of remoteness are made as rapidly or nearly as rapidly as are those of zero remoteness; and (3) associations from outside the lists are made more slowly than are any others.

The recall of items remote from the stimulus-word in both the forward and backward directions is clearly a fact under these conditions. The question may be raised, however, whether the present results may not be a function of instructions self-imposed by the *Ss*. Had they set themselves to give any but the next following item, something like these results might have been obtained. It does not seem likely that self-imposed instructions have produced the data. There is no known reason why the *Ss* should have

sought so uniformly to avoid sequent members, nor is it easy to understand why they should have done so at all since their most reasonable assumption would be that the association-test was another measure of learning and that the next following word was the one wanted. The fact, moreover, that the remote terms in either direction were given with a rapidity which is of the same order as the speed of the contiguous terms argues against such an interpretation of the results. Even if we assume, however, that the *Ss* did instruct themselves in this way, the fact that they were able to give remote terms so rapidly can only be taken to mean that remote associations had been formed.[17]

One may inquire, also, why, if the numbers of remote associations yielded by the present methods and conditions are valid, other experiments, such as those by the method of the derived list, have not found more extensive remote associations. The reason may well lie in the fact that the method of the derived list inevitably involves unknown amounts of associative and perhaps of other inhibition which may have masked the appearance of the remote associations. This method requires that the items be organized into a new sequence in which old associations either with position or with adjacent items must be broken and new ones formed. The association-method of recall requires no reorganization of the terms.

The remote associations measured by the association-method are not necessarily, moreover, exactly the same phenomena which the other methods measure. Remote associations measured by the method of the derived list, in terms of anticipatory errors, and by the association-method have been arrived at by very different operations and under different conditions. The results may be grouped in the same general class, but we can scarcely expect them to be identical.

If we grant the validity of the remote recalls in either direction, the question at once arises whether the connections between the remote terms were formed indirectly or as a result of direct practice whereby the two terms, though spatially remote, were reacted to by the *S* in immediate

[17] It should be noted in passing that the mean association-times of these experiments are very much shorter than those found by Wohlgemuth, *op. cit.* This fact may be ascribed to differences in experimental conditions, in particular to the instructions given to the *Ss*. His *Ss* were instructed to recall an item immediately preceding or succeeding the stimulus-item or at least an item from a position next but one. In the present experiments, on the other hand, our *Ss* were told to give the first word which came whether it was in the list or not. This permitted them to give words of any degree of remoteness within the list and to give words from outside the list as well. The wider range of choice would be expected to be accompanied by shorter association-times. The implications of the present data are concordant with those of Wohlgemuth.

sequence. The exposure interval was long enough, it might be insisted, to permit the *S* to recall other adjectives than the one presented while the latter was in the exposure-window and this recall would constitute a practicing of the two together. The subsequent recall of the represented word by the presented one, or vice versa, would be a case only of the recall of directly associated terms. Such recall during learning of other than immediately sequential terms does undoubtedly occur. But how could it have occurred in the first instance if remote associations had not already been established? Such recall might strengthen connections already made, but it depends upon the connections for its own occurrence. It is, moreover, very doubtful if *Ss* recall the non-presented terms as frequently or as uniformly from different parts of the list as would be necessary for this interpretation to fit the present results. *S* is forced to react to the forward moving sequence of items and his representative excursions must be incidental. Subtle as are representative processes, it is difficult to see how they could, under these conditions, even granting that the connections must not have been made prior to their entrance, serve to form so many forward and backward associations to a point at which they can function with a rapidity equal to or greater than that of the immediately sequent terms in the forward direction and without regard to degree of remoteness. It is particularly difficult to make this explanation fit the results of Experiment IV where the *Ss* found it very difficult to anticipate within the 1.8 sec. exposure-time and where there was, therefore, little opportunity to think backward and forward in the list. An explanation which goes beyond direct practice in the connection of spatially remote members is required.

We should, before considering possible theoretical interpretations, call attention to the relatively large percentages of associates from outside the list of which the stimulus-word was a member. This means that the learning of a list of adjectives to a criterion of 3 perfect trials does not uniformly establish intra-serial connections to a degree which causes them immediately after learning to be prepotent over competing responses. Recency and the other conditions of associative formation which have been active just before do not necessarily overweigh previously established connections.

If the remote backward and forward associations of these experiments are interpreted to be at least in considerable part valid cases of connection between temporally remote items, the further problem of the mode in which such remote association takes place is at once presented. A com-

plete and critical examination of the possibilities would lead one at once to the realm of theories of memory and of learning and thus to lengths beyond the scope of the present paper. On the negative side, however, these results cannot be interpreted on a simple chain-theory of serial learning. Both remote intra-serial associations and connections which lead outside the series are too numerous and too strong for that.

On the positive side, any theory will need to consider the influence of the change in set which accompanies the methodological shift from learning by the anticipation method to recall by the association-method where the direct sequence is not necessary. It should take account, also, of the fact that by the association-method there is a breaking up of sequence and of the antecedent context of each word since the stimulus-word is not now aroused by the preceding word and recalled by the *S*, but is spoken by *E*. It becomes, thus, a different stimulus-word.

Three theoretical possibilities may be stated. (1) The zero backward and all of the remote associations may be mediated by the connection of each item with some feature or features of the experimental situation. (2) There may be intra-serial mediating associations. (3) The instructions under this method may establish a set or determination which in some way operates independently of forward serial connections.[18] These three possibilities are not mutually exclusive and all may be involved. They could be elaborated at length, but theorizing may more profitably wait until we have a more detailed knowledge of the conditions which determine the phenomenon. The attempt to account for the fact that the frequency of remote associations does not vary inversely as degree of remoteness and that association times do not vary directly as remoteness, as some of the earlier work would lead one to expect, should wait upon further work by the present method and upon the formulation of an adequate theory of remote association.

Conclusions

The results of the separate experiments have already been stated in detail and it is necessary to give only the general conclusions here. The

[18] It may be, for example, that the change of set releases the remote associations from the inhibitions imposed by the set to repeat in the forward order. It might be also, that the association-method, which does not present the stimulus-words in their original serial order and according to which the intervening items between stimulus-word and response-word do not appear, acts in other ways to release remote associations from inhibition. Cf., in this connection, W. M. Lepley, A theory of serial learning and forgetting based on conditioned reflex principles, *Psychol. Rev.*, 39, 1932, 279-288; and C. L. Hull, The concept of the habit-family hierarchy and maze learning, *ibid.*, 41, 1934, 33-52, 134-152.

major conclusions refer to the results when scoring is on the assumption of linearity.

(1) The four experiments agree in finding a large number of recalls of items removed from the stimulus-word by one or more intervening serial positions. These number over half of all recalls in the forward direction in 3 of the 4 experiments and are distributed irregularly over the possible degrees of remoteness. There is no clear evidence of a correlation between frequency of remote association and degree of remoteness. They are recalled with a speed which is at least of the same order as that of the associations of zero degree of remoteness.

(2) There is also a considerable number of remote associations in the backward direction, but this is always relatively less than the number in the forward direction. Well over half of all backward associations are at least one term remote and they are given with a rapidity which is not greatly different from that of the associations of zero remoteness or from that of remote associations in the forward direction. There is a slight tendency for the frequency of remote backward associations to become less as degree of remoteness increases, but no definite conclusion on this point is possible.

(3) Varying percentages of associations come from outside the list of which the stimulus-word is a member. These associations have a higher mean association-time than do any of the others.

(4) The results obtained are not in their major characteristics a function of the particular conditions existing in any one of the four experiments. The same general conclusions emerge whether adjectives or Witmer consonants are the material, whether the list of adjectives is 10 or 16 items in length, and whether the test for remote association is made immediately after continuous learning to a criterion or after relearning following an interval since an earlier partial learning. The chief difference between the experiments is in the number of extra-serial recalls, which appear in larger percentages in the three other experiments than in Experiment I.

(5) When the data are scored on the assumption that the lists are learned as circular reaction-series, results of the same implication appear. There is somewhat more evidence for a decrease in frequency of remote associations with increasing degree of remoteness, but no evidence for a correlation of any great size between frequency of remote backward associations and degree of remoteness.

(6) These results uniformly establish the following major facts within

the conditions and operations used: (a) The existence of remote forward associations in serially learned lists; (b) the existence of backward association, both zero and remote, in serial lists learned in the forward direction; and (c) that forward and backward associations, whether of zero or of positive degrees of remoteness, do not differ greatly or consistently from each other in strength as measured by association-time. The remote associations in both directions and those in the backward direction at zero degree are interpreted, for reasons given in the discussion, as being temporally remote as well as spatially.

EXPERIMENTAL STUDIES OF THE ELEMENTS OF EXPRESSION IN MUSIC

By KATE HEVNER, University of Minnesota

The experiments to be described here deal with the problem of meaning in music, or, more explicitly, with the affective value and the expressiveness of music. In some earlier experiments the author had developed an objective and quantitative method for isolating certain variables in music structure and measuring their expressiveness, and by application of the method had investigated the affective value of the major and minor modes.[1] It is the object of the present study to improve the processes for treating the results from the method, and to discover the effect of three other variables, the rising and falling of the melodic line, the firm or flowing motion in the rhythm, and the simplicity or complexity of the harmony.

Obviously, in undertaking such a study, it is assumed that there is an expressiveness in music which is so generally and consistently associated with it that it will lend itself to experimental variation and control. We find it necessary to refer to this statement as an assumption because of the fact that in the history of musical culture, there has been a lengthy controversy over the problem of expressiveness in music, and much difference of opinion as to whether music can be made to express definite emotions and concepts and sentiments, or whether it is merely suggestive and moving in a very general way and to be interpreted quite variously by the different listeners who are attending to it. On the one hand, we find the very important and successful use of musical expression in the modern opera, the tone poems, and other forms of descriptive music which have met with great popular success. The volumes upon volumes of program notes also testify to the popularity of the assumption with concert audiences. Likewise, in experimental psychology we have evidence that at least the general mood or affective state suggested by any one composition is fairly constant and universal.[2] The introspective reports of individuals who try to analyze their experience in listening also show that this interpretative activity is of very frequent occurrence,[3] and furthermore that it is often the means by which apprehension or appreciation of a more difficult type is developed. On the other hand, it is a simple matter to demonstrate that listeners often disagree markedly in their specific interpretations of music, and some of the differences of opinion of eminent critics on important symphonies have become world-famous. It is certainly clear at least, that there are different kinds of music, formal and descriptive, absolute and pragmatic, pure music and music full of symbolism, and also that there are different kinds of listeners to music. In short, there are many ways of enjoying music, and many

* Accepted for publication July 1, 1934.

[1] K. Hevner, The affective character of the major and minor modes in music, this JOURNAL, 47, 1935, 103-118.

[2] Max Schoen, editor, *The Effects of Music*, chaps. VII, VIII.

[3] Max Schoen, The aesthetic attitude in music, *Psychol. Monog.*, 39, 1928, (no. 178), 162-183.

types of music to enjoy, and of exquisitely delightful music of all varieties, and enthusiastic listeners of all kinds, there has never been any lack.

Since there has been so much confusion and disagreement attending the problem of meaning in music, it will be necessary to state clearly at the outset what phase of this complex question we are studying, and what further assumptions and theories we endorse. The limitations of expressiveness, its psychological explanation, its use in understanding and appreciating music, and the particular kind of symbolism which it employs, we shall discuss at length elsewhere;[4] for the present we shall sketch the underlying theory and method of our experiments as briefly as possible.

The presence or absence of meaning, and the particular quality of the meaning in any music is dependent on a number of factors: the form and structure of the music itself, the attitude of the listeners, their previous experience, their training, talent and temperament, and their momentary mood and physiological condition. A certain flexibility in expressiveness is therefore postulated, and experimental studies will be the means of determining the nature and limits of this flexibility.

The suggestiveness of music does not depend directly on the amount of formal training which the listener has had, but is apparent both to the trained and untrained. Previous experiments have shown that the perception of meaning is aided and facilitated to some extent by musical training, experience, and interest, but that the interpretations are affected by these factors much less than popular opinion has supposed.

Studies of a special group of students with no background or training in music have shown that they differentiate the mood effects of the major and minor mode quite distinctly and readily. The measurements show that the typical or median student in this special group has never had any private music lessons of any sort, does not play any musical instrument, has no musicians in his family and never attends operas or concerts. He can read music a little for singing, has sung two years in a high school chorus and has a piano and a radio in his home; and he is almost as suggestible to the mood effects of music as his highly trained classmate.[5]

The *Os* (450 in number) have, in the experiments to be described below, been measured for their musical interest, training, attitude, and appreciation, and they represent a very wide range on all of these scales, with the large majorities falling into the untrained and untalented groups. They include a few students who are preparing for a professional career in music (usually teaching), and a few who have an aversion to music and avoid all contact with it; but for the most part they claim only a casual interest in music, have had a lesson or two, exhibit some little preference, have no knowledge of the actual form and structure of musical compositions,

[4] K. Hevner, Expression in music: A discussion of experimental studies and theories, *Psychol. Rev.*, 42, 1935, 186-204.

[5] Hevner, *op. cit.*, this JOURNAL, 47, 1935, 115 ff.

and are strikingly lacking in a vocabulary with which to describe what they hear and with which to identify musical sounds.

The symbolism of music is obviously not as exact or as specific as verbal symbolism, but we begin with the assumption that it is not chaotic, and not entirely without system. On the contrary, we believe that it may be possible to discover the presence of definite principles together with a certain orderliness in the relation between musical structures and the various emotions and sentiments.

In our experimental search for those elements of musical structure which must be at the basis of any such orderly system, there are certain details in methodology which must receive careful consideration. Since we are looking for elements of *music* we must be sure that the material provided for observation represents real *music* and not merely *elements* trimmed down for experimental purposes to such an extent that all the *music* has been left out. The outline of a rhythmic pattern, to choose one of the possible elements, tapped out with a hard wooden stylus, is but the bare skeleton of a rhythm, rattling its dry bones in vast emptinesses, and far different from the living, throbbing rhythm that pulsates through the whole body of a musical composition. The isolation of the variable to be studied must not be accomplished by an actual isolation of the one particular element from all the relationships which make it musical, but by some better method; by the study of various compositions in which this one element predominates, or by arranging several versions of one musical composition which differ from each other only with respect to the element under consideration.

We must also recognize the importance of the *cumulative* effect which the succession of stimuli makes possible, and on which the effectiveness of some of the elements largely depends. The repetition of certain stimuli and their natural support by other qualities in the body of the music is necessary in order that the human organism may perceive them. The sadness which is so strongly associated with the minor does not reside in the minor chord alone, but only in music written in the minor mode, with all the accumulation and interrelation of stimuli which the use of that mode implies. Again, music is not a spatial, but a temporal art, and distinguished from the visual arts by its dependence on temporal rather than spatial relationships. In choosing the elements which carry the meaning of music, and which will form the basis of a system of musical symbolism we must be careful, therefore, to choose those elements which involve successive moments of time, such as rhythm and tempo, rather than intervals and chords. It is the masses of harmony, the resolutions and progressions, to which meaning attaches rather than to the triads themselves, and it is the twists and turns of the melody around its keynote, the running of the melodic line, rather than certain momentary skips or intervals, which carry the suggestiveness and meaning of the music.

In securing the responses of the individual listeners, there are also many sources of error to be avoided. The responses should be as objective and simple as possible in order that they may be summarized quickly and accurately by *E*, and treated quantitatively. But *O* must not be forced into a judgment of the meaning when he does not spontaneously respond in that particular way, and he must be provided with a wide selection of responses, so that he may be exactly suited in his final choice. From the point of view of both *O* and *E*, a most satisfactory method of recording responses is by means of a list of 66 adjectives, arranged in 14 groups,[6] from which

[6] The list of adjectives is given in an earlier study, Hevner, *op. cit.*, this JOURNAL, 47, 1935, 109.

the listener checks all the adjectives which he finds appropriate to the music. This makes the business of choosing responses easy and convenient for *O*, and allows an objective and quantitative treatment by *E*. From the original response sheets of the listeners, each of whom had a fresh copy of the adjective list for every different musical composition, it was a simple matter to tabulate all the votes for any one adjective from all the *O*s for each composition alone, or for all the compositions which are alike in certain respects. Then, by comparing the numbers of votes for different adjectives, the meanings or affective characteristics of the compositions could be ascertained.

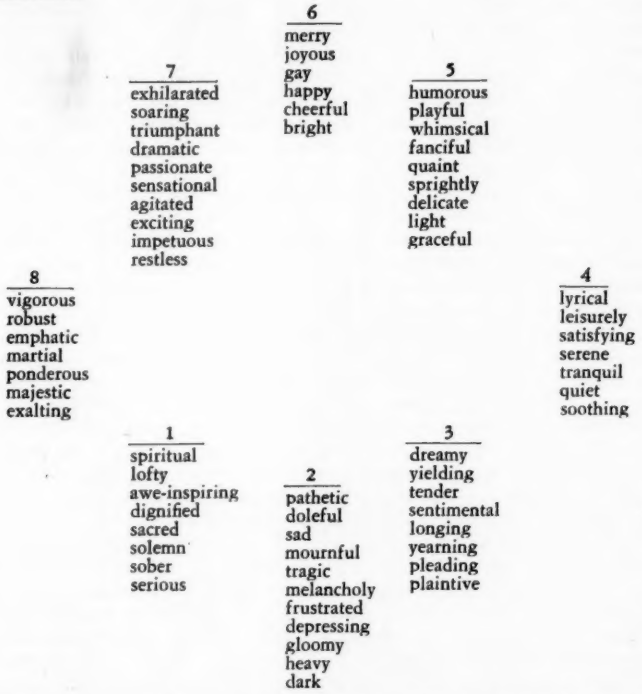


FIG. 1. SHOWING THE ADJECTIVES FROM LIST I ARRANGED IN 8 RELATED GROUPS

But a still further advantage has been gained for the use of this adjective-list method by a process of treating the data after they have been tabulated, a process of combining the individual adjectives into homogeneous groups, forming a more or less continuous scale. This process makes allowance for the errors arising either from the small disagreements on the part of the listeners about the exact definitions of words (which are important perhaps to the *O*s but certainly not to *E*), and from the effects of the momentary mood and physiological condition of the different *O*s which may modify slightly the mood-effect produced by the music. It also simplifies

tremendously the task of summarizing the data for comparing the results of different musical compositions, because of the fact that the adjectives are not treated as 66 individual entities, but are arranged into groups so that all those with the same feeling-tone may be considered together.

This grouping was accomplished in consultation with several *O*s who attempted to arrange their groups of adjectives around an imaginary circle in such a way that the adjectives in one group would be closely related and compatible with each other; that any two adjacent groups should have some characteristics in common, and that the groups at the extremities of any diameter of the circle should be as unlike each other as possible. The arrangement in Fig. 1 seemed the most satisfactory. This arrangement is not the only possible grouping along such a scale, and does not perfectly satisfy all the demands of the opposing groups. It has value merely in simplifying and summarizing the results of these studies. The adjacent groups have certain characteristics in common, and the transitions between them are made without abrupt or awkward changes in feeling tone. A complete circuit of the circle carries one quite smoothly over the range of the most common affective experiences. In studying the arrangement in the circle, we may note several details which are consistent with our usual concepts of affective states. Muscular tension is probably most strongly associated with the vigorous-robust exciting-impetuous groups, and least with the groups directly opposite, the sentimental-yearning and quiet-serene groups. Likewise, activity is associated with the four groups beginning with vigorous-robust through exhilarated-restless, merry-bright, and humorous-graceful, but quiescence is associated with the other four groups, from lyrical-soothing to spiritual-serious.

For our first experiments, we used phonograph recordings of 5 different musical compositions as follows: Debussy, *Reflections on the Water,* a piano solo played by Paderewski; Mendelssohn, *Midsummer Night's Dream, Scherzo,* played by the Philharmonic Symphony Orchestra of New York under Toscanini; Paganini, *Étude in E Flat Major,* a Liszt-Busoni arrangement, played as a piano solo by Vladimir Horowitz; Tschaikowsky, *Symphony No. 6 in B minor,* the second part of the first movement, played by a European orchestra conducted by Albert Gates; Wagner, *Lohengrin,* Prelude to Act III, played by the Boston Symphony Orchestra. These were all Victor records, and the first four were orthophonic recordings.

These 5 records were played to a group of 52 *O*s, who were directed to check from the printed list all the adjectives which seemed appropriate for describing the music and to check as few or as many words as they liked. Of course, no hint of the name or nature of the music was given by *E*, and by inquiry it was found that one person in the group could be expected to identify the music exactly, and one person more indefinitely as "something from Lohengrin," or "a Beethoven Symphony." The *O*s were students in the Elementary Laboratory course in psychology; the majority were sophomores, and about one-third of them were men. When the votes for each adjective were tabulated, it was found that in the case of each of the 5 compositions there were some *O*s who could find no adjectives appropriate, or at the most only one or two; but most of the *O*s checked 6 or more adjectives, and found it an easy and agreeable task.

The graphs in Fig. 2 show the results of this experiment. In order to plot the diagram for, let us say, the Tschaikowsky Symphony, all the adjectives used from one group were added, and plotted on a base-line, together with total number of adjectives used from all the other groups. In other words, every group of adjectives is

treated as a whole, as representing one kind of feeling tone. For convenience in plotting these graphs, the groups are arranged, not around a circle as in Fig. 1, but

FIG. 2. NUMBER OF *Os* SELECTING ADJECTIVES FROM EACH GROUP

on a straight line, with the dignified-solemn group at the extreme left, and the robust-vigorous group at the extreme right. The reader must be warned that such an arrangement does not imply that the groups actually represent a linear scale with numerical

units. We have chosen these groupings quite arbitrarily, (1) because it was the opinion of the persons who first studied the adjectives that this was the best arrangement to bring out the logical and physiological realtionships between the groups, and (2) because empirical data, collected with these groupings on various kinds of aesthetic or affective material, poetry, music, and paintings, show a satisfactory 'patterning' which very much simplifies the work of presenting and summarizing results. The graphs in Fig. 2 represent the total number of adjectives used from each group, divided by the number of *Os* making the selection. This process of division is always necessary in order to make these graphs comparable to the graphs from later experiments with different groups of *Os* ranging in number from 20 to 100 plus. The graphs are always made in percentages, consequently the results from a group of 40 *Os* may be compared directly with those from a group of 65, etc.

In studying the graphs, the consistency and uniformity of the interpretation of the *Os* is immediately apparent. The reader must remember that adjective-groups 1 and 8 do not represent the extremes of a scale, but are as closely related as any other two adjacent groups on the scale, and that the pattern of the Lohengrin music shows as much consistency, with its concentration of votes for agitated-impetuous and vigorous-robust, its absence of votes in the central area, and its upthrust on the dignified-majestic group, as does the pattern for the Mendelssohn, where the concentration is more readily apparent to the eye. The diagrams are valuable, not only for the consistency and uniformity which they show, but also for the appropriateness of the interpretation in consideration of the composer's intention. Mendelssohn's *Scherzo,* to be played after the first act of Shakespeare's *Midsummer Night's Dream,* is outstandingly exciting-impetuous and playful-graceful, and there is a total lack of such affective states as serenity, sadness, dignity, and solemnity. The Prelude to the third act of *Lohengrin* stands out as impetuous-exciting, robust-vigorous, and it is even more interesting to find on the examination of our original tabulations that the heavy voting for this composition falls on the adjectives exhilarated, triumphant, majestic, and exalting. The Paganini *Étude* (a composition characterized by its emphatic rhythms and rapid tripping arpeggios, and presented by Horowitz with great brilliance and precision) lacks entirely the feeling-tone of sadness, dignity, or serenity, and suggests playfulness, gayety, excitement, and vigor. These interpretations are all the more significant because they have been made by listeners distinctly of the untrained type.

The Debussy composition, *Reflections on the Water,* is of especial interest because of the apparent uncertainty of the listeners, and the discrepancies which appear in the data. The adjectives used are scattered among the 8 groups with the greatest numbers from two groups in diametrical opposition to each other, viz. dreamy-sentimental and exciting-impetuous. There are two reasons which might account for these differences in interpretations. The first is the nature of the subject with which the composer is dealing. Reflections on the water are likely to be at one time clear, still, and distinct, and at another time distorted, restless, fitful, and changing, and the composer may have intended several different interpretations to be possible from his music. From this point of view the votes of our listeners may stamp this composition as a particularly successful bit of creative work. The second reason, which is important for many other compositions as well as this particular one of Debussy's, is the presence of several well-defined sections in the musical structure, sometimes quite different from each other. One of the commonest patterns for musical composition consists in a

division of the material into three parts, with the first and third similar or even identical, and the second a contrasting section, only remotely related to the preceding and succeeding parts.

In order to separate the various parts of the composition and trace the source of the votes for the different adjective groups, the experiments described above were repeated with a new group of *O*s, and with slightly different instructions. The new *O*s began to check their adjectives as soon as the suggestiveness of the music became apparent, by placing the numeral '1' before each choice. At a given signal from *E* which coincided with the beginning of a new section of the music, the *O*s indicated their choices by the numeral '2'; and so on for a third (a fourth or fifth) section, they changed to numeral '3,' etc. Tabulations were made for the three different sections of the piece, and the graphs which show these results are given in Fig. 3. The reactions occasioned by the first part of the composition appear in diagonals, for the second part in solid black, and for the third part in white areas. The separations for the *Reflections on the Water* were made at bars 22 and 75 and the sources of the different affective patterns are very clearly shown. The first and third sections are quite similar in character: dreamy-sentimental, serene, calm and slightly graceful, and playful. In sharp contrast, the middle section is very playful, rather merry, and extremely restless, exciting, and agitated.

With the two experiments carried out under two sets of directions, it would not be logical to expect that combining all the adjectives numbered 1, 2, and 3 in the second experiment would give the same pattern when plotted as the graph for the same music from the first experiment (Fig. 2). An inspection of Fig. 3 shows that if such a procedure were followed, there would be a great preponderance of votes for certain adjective groups, especially Groups 3 and 4, because the recurrence of the first section at the end of the composition makes it necessary for the listeners to vote for those adjectives twice. A comparison of the graphs in Figs. 2 and 3 does reveal, however, the fundamental agreement in the interpretations of the two groups, while the differences due to the slightly different experimental conditions are distinctly brought out. It is quite likely that the *O*s maintained a much higher degree of attention to the music in the second experiment than in the first.

In Fig. 3, results are given also for several other compositions. The Tschaikowsky *Symphony* is of especial interest, for the excitement and agitation are shown to originate in the second section which is higher, faster, and thinner than the first, and from the third section, which is a repetition of the first with much more volume and with a more elaborate orchestration, including a prominent glissando effect in the violins.

Two of Beethoven's ~~*Sonatas*~~ were also used in this part of the study, the *Seventh, in A major*, third movement, first half, and the *Fifth, in C minor*, first movement, first half,[7] and the results are shown in Fig. 3. In both of these compositions the first and third parts are alike and the second part is different in tempo, rhythm, theme, and orchestration. The graphs of the mood-effects bring out these factors very clearly, especially in the excerpt from the *Fifth Symphony*, where the first and third sections were identical in every way. The more complicated graph for the Mendels-

[7] The music is played by Felix Weingartner and symphony orchestra, Columbia record. In the Fifth Symphony the music is divided at the 71st and 125th bars; in the Seventh at the 145th and 237th.

sohn *Scherzo* is not included, although its five different parts showed equally interesting and consistent results.

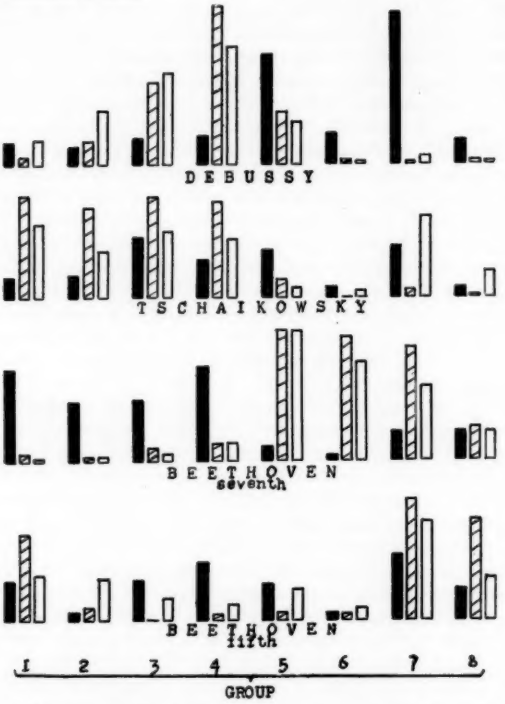
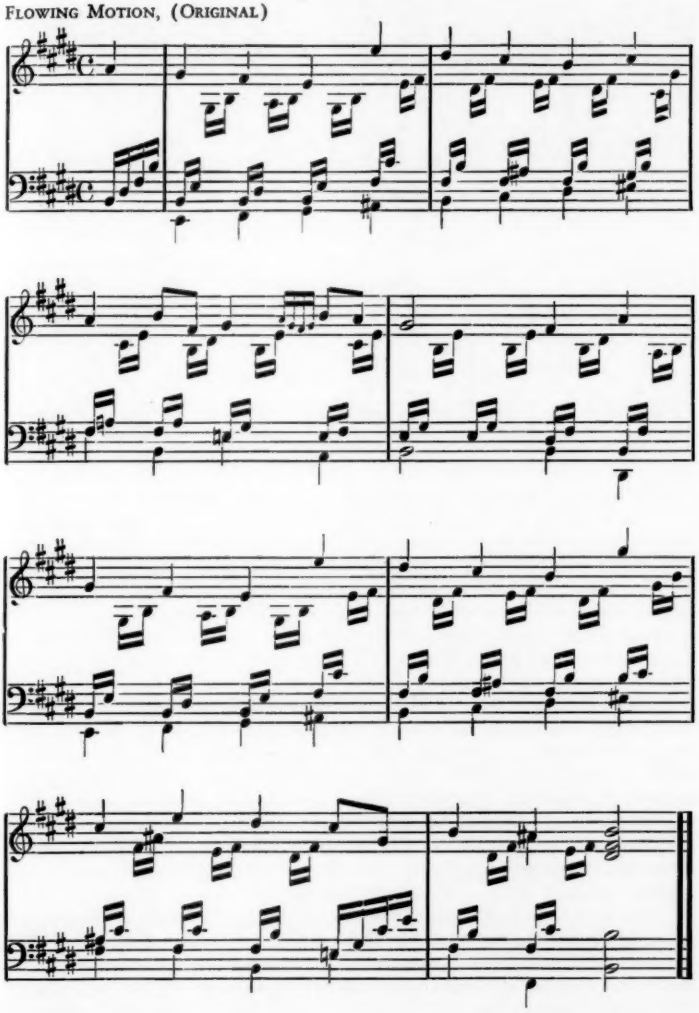


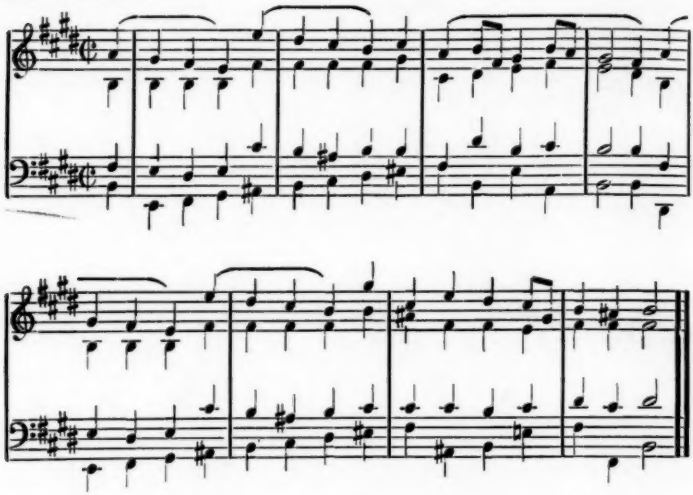
FIG. 3. NUMBER OF *Os* SELECTING ADJECTIVES FROM EACH GROUP

Since the adjective-group method was shown to be successful in handling the data from these longer compositions, the same process has been applied also to the study of the affective character of certain *elements* of musical structure. The method of isolating the element to be studied follows the method described in an earlier experiment, on the affective qualities of the major and minor modes.[8] Two versions of the same composition were presented to two different audiences who were asked to express their attitude toward the music by checking adjectives. The music in the two versions is arranged in such a way that the difference between them lies only in the one element to be studied, and the listeners indicate their interpretation of the music

[8] For a complete discussion of the advantages and disadvantages of composing experimental music in this way *cf.* the author's previously cited article. The music must be carefully chosen, and rewritten by the most expert musicians. For invaluable assistance in this work the author is indebted to Miss Hariett Johnson, pupil of Rubin Goldmark, Lecturer for The Layman's Music, Inc., New York; and to Mr. John Verall of the Music Department, Hamline University, St. Paul.

MENDELSSOHN SONG WITHOUT WORDS
(Descending melody)

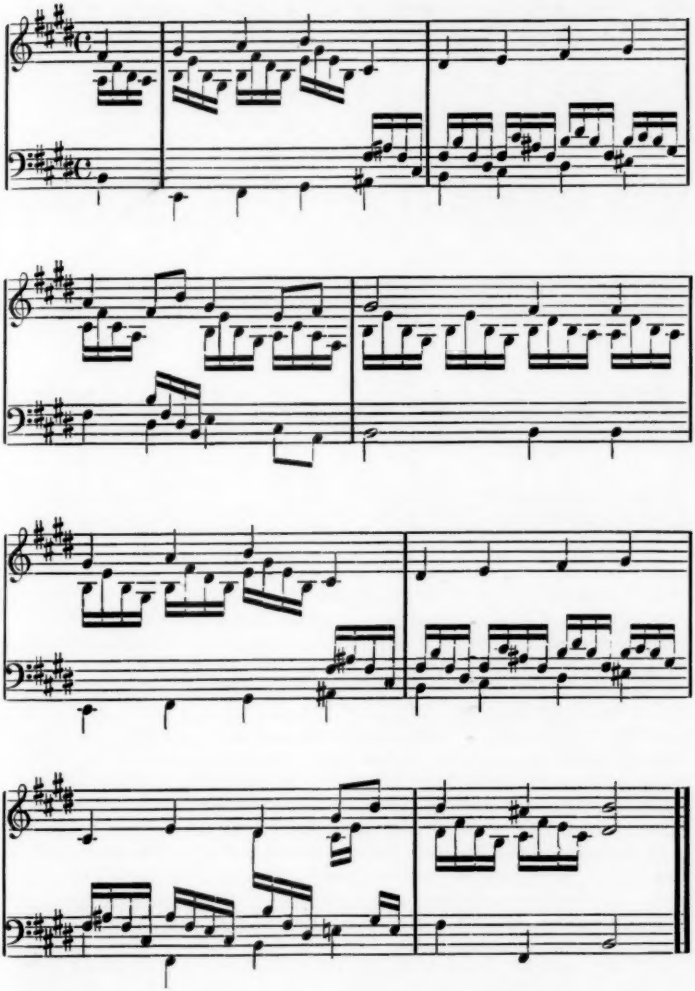
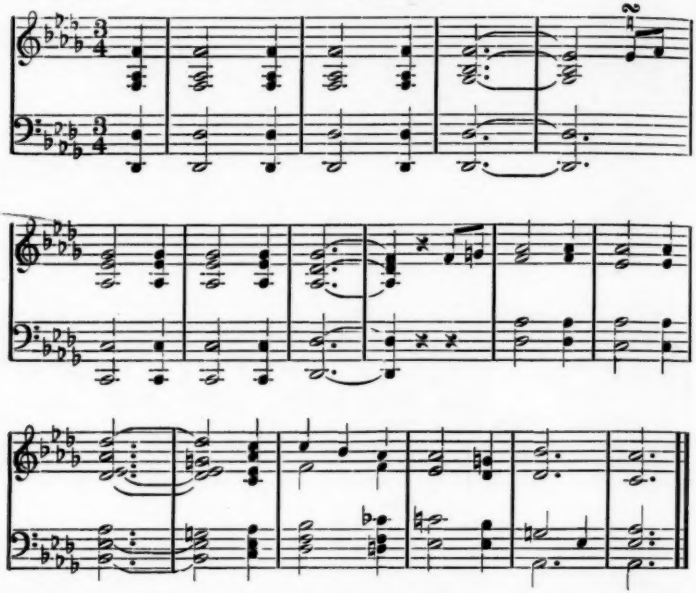
FLOWING MOTION, (ORIGINAL)

FIRM RHYTHM



by checking all the adjectives, many or few, which they find appropriate. In no case did the same group of listeners hear both versions of the same composition. The program of music for any one group was arranged to include 10 different compositions, with a wide variety of rhythms, melodies, and harmonies. They were played by the same pianist, who had practiced under observation until the two versions of each composition were executed at the same tempo and intensity and with the same interpretative effects. The objective records of the critic observer on prepared schedule blanks show that her interpretations of the music exhibit an extremely high degree of similarity for the various performances.

In the first group of 6 compositions the element changed was the *rhythm or motion* of the music. The motion was changed from a firm beat with a full chord on every beat as in a chorale or hymn tune, to a much more smooth and flowing motion in which the supporting chords were broken up and spread evenly throughout the measure. A part of the accompanying harmony occurred not only on each beat but on every half beat with the result that the motion of the flowing version was more continuous and less definitely accented than that of the firm version. As an example of the type of change, we will reproduce the music scores for twelve measures of a Mendelssohn *Song Without Words,* Op. 19, No. 1.

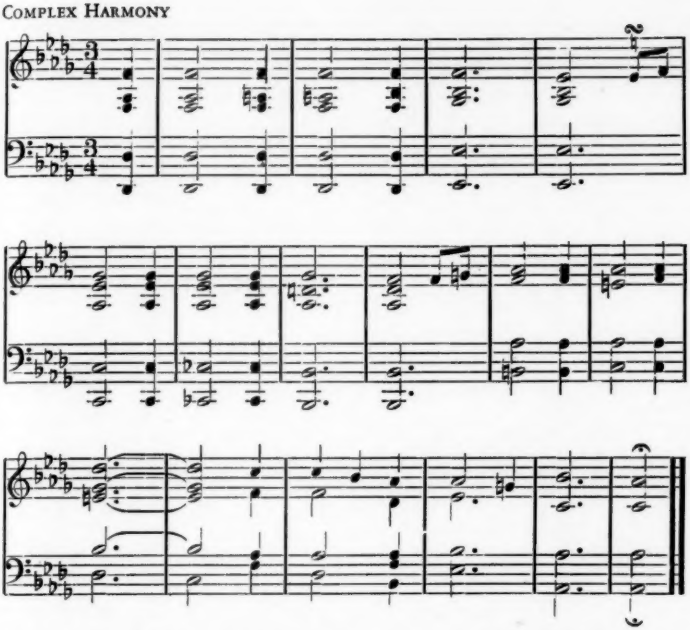
In the second group of 9 compositions the variable to be studied was the *rising and falling of the melodic line.* In preparing the experimental material, the author chose a number of short compositions representing a variety of styles and characters, sad, gay, serene, exciting, etc., and a variety of melodic lines. In the Beethoven Op. 13, for example, there is a long and vigorous rise in the melody through two octaves. In the Mendelssohn Op. 19, the melody opens with three slow climbing phrases. In the Schumann Op. 15, No. 5, the melody is strongly descending in character, a short descending motif carried down through 4 octaves. In other compositions the original

MENDELSSOHN SONG WITHOUT WORDS
(Ascending melody)

BEETHOVEN, OP. 10. SIMPLE HARMONY, (ORIGINAL)

melody had but a very slight rise or fall, and in others it was practically stationary. In each of the compositions the original melody was inverted. The inversion was done as exactly as possible; the skips were of the same size and location, but in the opposite direction from those of the original melody. As an example of the treatment, we quote here another version of the Mendelssohn *Song Without Words,* Op. 19, No. 1, an ascending version, which the reader may compare with the descending melody quoted on page 255.

Of the compositions on which this process was first tried out, many melodies were found to be unsuitable, just as many melodies cannot be used as fugue themes, because they do not lend themselves to the intricate manipulation necessary in fugue composition. In choosing and rewriting the music for these studies the most important consideration was that of making the two versions different in one factor only, the ascending or descending character of the melodic line. The many difficulties that arose in performing this task were met in different ways, as follows.

(1) The new melody resulting from the inversion did not always sound sensible or logical, and was often unmusical, and even unpleasant. Many compositions were tried and discarded for this reason.

(2) The harmony demanded by the new melody was not always the same as the harmony of the original and hence, certain harmonic changes were effected in the new version. In making these changes, the general character and style of the harmony was always maintained, although the dominant triad sometimes became a dominant seventh, the VI chord was sometimes changed to the II, etc. Compromises

COMPLEX HARMONY

had to be made between the demands of the harmony and the exactness of the inversion of the melody.

(3) It was sometimes difficult to keep the relation of the melody to the tonic or keynote exactly parallel throughout the length of the two versions. The keynote is the point of reference for all parts of the melody and returns to it are important because they give a feeling of satisfaction and finality which helps to establish the form of the composition and give it balance and unity. The parallelism was always strictly maintained, however, at the ends of the phrases, since they mark off the melody into definite parts and determine the kind of symmetry it shall have. It is understood then that the inversions of the melodies were never perfect, but that they were as exact as these other considerations could allow.[9]

In the third group of 9 compositions the modifications were made in the harmonic structure of the music, changing it from simple, consonant harmonies, resolved in a very smooth and satisfying manner, to complex, dissonant harmonies, more harsh and rough, and not always reaching a satisfactory resolution. The versions which we call complex have augmented and diminished intervals and modern effects. It will be surmised from the discussion of the difficulties encountered in making melodic

[9] The difficulties encountered in inverting melodies were so great that this part of the experiment was entirely revised and repeated, with new music and new *O*s. The music for the second experiment was much more satisfactory to the music critics, but the results from the two series are very little different, and therefore only the second experiment is reported here.

TABLE I

SHOWING THE VOTES FOR EACH ADJECTIVE GROUP, FOR 6 COMPOSITIONS IN FIRM RHYTHM AND IN FLOWING RHYTHM

	Adj.-Group	Beethoven, Op. 13	Schumann, Op. 15 No. 5	Mendelssohn, Op. 19 No. 1	Schubert, *Sonata*, No. 120	Prokofieff, Op. 12	Schumann, Op. 18	Total	$\frac{D}{PE_D}$
Firm	1	11	79	126	27	3	36	282	17.7
Flowing		14	4	8	6	17	10	59	
Firm	2	30	32	20	9	10	13	114	2.7
Flowing		11	7	10	0	26	14	68	
Firm	3	5	0	37	17	15	14	88	9.5
Flowing		13	12	57	33	39	32	186	
Firm	4	2	4	73	72	21	18	190	1.9
Flowing		6	33	64	51	35	27	216	
Firm	5	5	2	14	42	94	21	178	7.7
Flowing		16	100	32	42	46	40	276	
Firm	6	11	3	1	45	50	13	123	10.4
Flowing		26	86	14	68	28	19	241	
Firm	7	143	68	4	11	13	33	272	1.7
Flowing		91	32	13	12	27	71	246	
Firm	8	68	116	8	22	4	21	239	10.4
Flowing		29	8	4	6	12	14	73	

TABLE II

SHOWING THE VOTES FOR EACH ADJECTIVE GROUP, FOR 9 COMPOSITIONS IN SIMPLE AND COMPLEX HARMONY

	Adj.-Group	Johnson, *Gigue*	Bach, *Chorale*	Gluck, *Andante*	Johnson, *Vigorous & Lively*	Verall, *Chorale*	Mendelssohn, Op. 64	Beethoven, *Sonata* Op. 10	Grieg, *Wedding Day*	Palestrina, *Choral*	Total	D / PED
Simple	1	0	75	9	2	149	4	141	11	74	465	3.0
Complex		0	59	34	3	82	39	83	48	50	398	
Simple	2	0	88	9	0	135	13	29	6	54	334	7.3
Complex		0	70	24	1	105	45	74	48	90	457	
Simple	3	0	49	84	0	6	143	31	39	26	378	3.7
Complex		1	4	83	0	19	99	27	32	48	313	
Simple	4	6	41	139	6	8	114	27	33	35	409	10.1
Complex		2	0	96	1	7	71	11	15	8	211	
Simple	5	126	3	122	91	0	19	2	31	1	395	12.0
Complex		91	0	48	32	0	9	2	2	2	186	
Simple	6	166	0	37	96	0	2	0	22	0	323	16.0
Complex		75	0	9	32	0	0	0	0	0	116	
Simple	7	39	0	1	88	2	1	4	14	1	150	14.5
Complex		114	14	4	194	10	8	42	7	7	400	
Simple	8	8	8	0	21	23	0	24	12	18	114	8.3
Complex		28	32	2	55	26	5	37	9	9	221	

TABLE III

SHOWING THE VOTES FOR EACH ADJECTIVE GROUP, FOR 9 COMPOSITIONS IN ASCENDING AND DESCENDING MELODY

	Adj.-Group	Beethoven, Op. 13	Gluck, *Caprice*	Grieg, *Wedding Day*	Verall, *Prelude*	Schumann, Op. 15 No. 5	Mendelssohn, Op. 19	Verall, *Nocturne*	Schumann, Op. 18	Mendelssohn, Op. 62	Total	D / PED
Ascending	1	5	56	27	15	31	5	5	2	27	173	4.5
Descending		21	8	12	15	21	9	15	0	14	115	
Ascending	2	2	10	25	28	5	6	7	0	27	110	.00
Descending		10	7	4	43	2	5	17	0	14	102	
Ascending	3	0	22	19	34	0	40	26	5	83	229	.20
Descending		2	32	30	24	0	44	52	0	27	211	
Ascending	4	1	28	14	20	0	66	17	15	55	216	3.8
Descending		1	40	17	20	0	77	63	8	23	249	
Ascending	5	0	53	14	10	0	17	31	80	4	209	2.6
Descending		0	20	9	42	0	34	65	46	11	227	
Ascending	6	2	11	8	1	6	5	8	72	1	114	2.5
Descending		4	4	6	9	9	14	15	42	4	107	
Ascending	7	65	1	10	0	78	0	2	20	3	179	6.6
Descending		178	0	0	11	42	2	1	16	1	251	
Ascending	8	38	3	12	1	130	1	0	0	1	186	1.00
Descending		115	0	3	2	61	2	1	0	0	184	

TABLE IV

SHOWING THE VOTES FOR EACH ADJECTIVE GROUP, FOR 10 COMPOSITIONS IN MAJOR AND MINOR MODE

	Adj.-Group	Schumann, *Folk Song*	Beethoven, *Minuet*	Bach, *Minuet*	Martini, *Gavotte*	Bach, *Musette*	Rameau, *Rigaudon*	Durand, *Chaconne*	Gluck, *Andante*	Arensky, *A Song*	Rameau, *Tambourine*	Total	$\frac{D}{PE_D}$
Major	1	20	20	14	8	13	10	1	95	209	5	395	4.0
Minor		43	74	20	28	46	21	4	108	144	23	511	
Major	2	7	20	4	0	7	7	0	38	163	5	251	20.0
Minor		67	95	17	30	111	56	3	122	224	23	748	
Major	3	33	68	10	13	43	16	4	75	105	8	375	12.0
Minor		107	140	32	68	111	42	9	125	84	29	747	
Major	4	53	99	49	45	82	33	30	131	127	27	676	3.4
Minor		52	124	60	40	71	29	22	94	57	25	574	
Major	5	152	104	225	210	139	108	287	47	8	245	1525	21.2
Minor		90	87	97	107	83	87	134	25	4	102	816	
Major	6	174	69	215	244	93	128	335	11	0	210	1479	24.2
Minor		64	21	67	93	25	61	155	4	0	114	604	
Major	7	42	27	102	59	17	47	84	16	11	115	519	0.00
Minor		43	24	66	22	35	95	96	32	7	125	545	
Major	8	16	6	32	31	5	34	30	15	12	57	238	0.00
Minor		20	11	23	12	9	41	22	23	18	64	243	

versions that the new harmonies sometimes wrought changes in the melodic line, so that it became a slightly different melody, even though the sequence of tones remained identically the same. To meet this argument, we can but repeat that we were aware of such dangers, and have guarded against them as far as possible in the rearrangements of our music. In order that the reader may have the opportunity of judging for himself in these matters we quote one typical example of our musical material, eight measures from Beethoven, Op. 10.

Table I shows the number of times each adjective group was selected for both the firm and flowing versions of the 6 different compositions. The total number of votes for all 6 compositions is also shown, and the ratio of the difference to its probable error for each group.[10] These ratios, shown in the last column of the table, are the most important figures for estimating the difference in the affective value or expressiveness of the two different rhythms, and therefore are expressed graphically in Fig. 4. Tables II and III present similar data for the rising and falling of the melodic line and for the simplicity and complexity of the harmony. Table IV presents data from an earlier experiment with the major and minor modes.

In Fig. 4, the significant data from all four tables are presented in graphic form. The octagonal figures represent the 8 different adjective groups designated here by one or two characteristic adjectives from each group. (The reader may refer to Fig. 1 for the full list of adjectives in each group.) There is a separate figure for each of the variables studied, modality, rhythm, harmony, and melody. The ratios of the differences to their probable errors (D/PE_D) are represented for each group in an appropriate symbol, so that the reader may see at a glance the importance of each element in differentiating the mood effects of the music.

It is apparent that the use of the major or minor mode is of the most clear-cut significance in the expression of four different mood effects, that the major mode is very strongly associated with happiness, gayety, playfulness and sprightliness, and the minor with sadness, and with sentimental yearning, tender effects. The expressiveness of modality, either major or minor, is more stable and more generally understood than that of any other element which we have studied. It is of equal interest to note also that the modality has no discoverable effect on the expressiveness of some of the affective states, on dignity, vigor, excitement, or calm. Those states can be produced equally well by means of either mode.

The firm or flowing motion in the rhythm divides the circle of adjectives exactly in two, and has its most important effects on a different axis of the circle, on the dignified and vigorous *vs.* the happy, playful, sprightly groups. These differences fall where musical tradition and custom would

[10] Calculated by Yule's formula and the Edgerton-Paterson tables, *J. Applied Psychol.*, 10, 1926, 378-391. The total number of adjectives used is taken as a percentage of the total number possible, *i.e.* the number of *O*s multiplied by the number of adjectives in the group, usually a total of between 3000 and 4000. The ratios would all be much larger save for this very conservative way of estimating the percentages, for use by every *O* of every adjective is not, after all, a real possibility.

lead us to expect them and on the whole seem quite logical and acceptable.

The differences brought about by simple *vs.* complex harmony, or we might say by consonance *vs.* dissonance or classic *vs.* modern harmony, are for the most part effective on the 5 adjective groups on the upper half

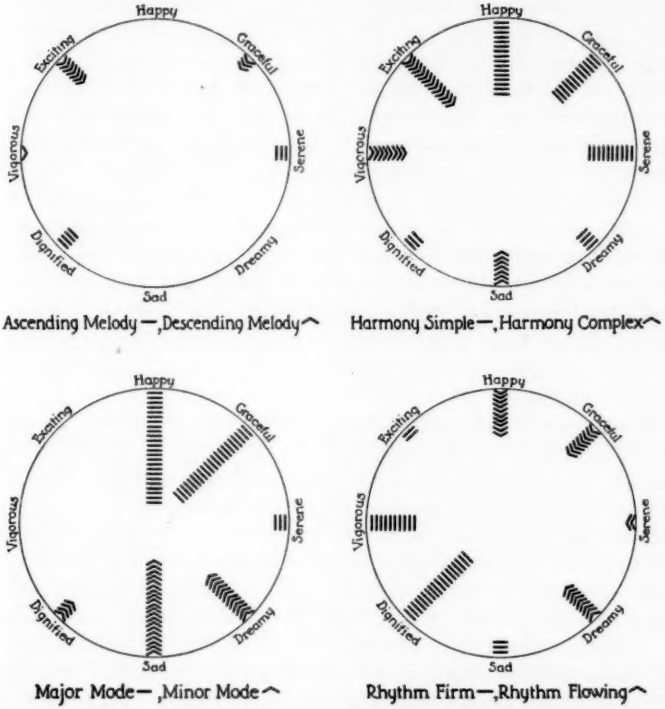


FIG. 4. SHOWING THE RELATIVE IMPORTANCE OF MODE, RHYTHM, HARMONY, AND MELODY ON THE AFFECTIVE VALUE OF MUSIC

of the circle. The simple consonant harmonies are happy, graceful, serene, and lyrical. The dissonant modern harmonies are exciting, agitating, vigorous, and somewhat sad. The differences are substantial, but not as clear-cut as those for the major and minor modes.

In comparison with the other three figures, the circle representing the effects of the rising and falling of the melodic line is particularly noteworthy for its lack of significant differences. The largest difference, indicating that the descending melodies are more exciting and impetuous,

is only 6 times its probable error. The importance of this finding is minimized by the difference likewise in favor of the descending versions, for the adjective group calm, lyrical, serene, etc. The descending versions give only one significant difference, on the dignified-serious group.

The first inference to be drawn from a study of these results is the relative importance of these four elements in expressing the affective tone of music. Apparently the melody, at least the rising or falling of the melody, is of much less importance than the harmony, or rhythm of the music, and the modality is most important of all. It must be remembered that the figures indicate the expressiveness of these elements only in relation to one another. Compositions written in the major mode are not always gay and playful and nothing else. They are only *more* gay and *more* playful than the same compositions played in the minor mode. Likewise, a composition with a firm rhythm will be much more serious, dignified and vigorous than the same composition with a flowing rhythm. But a composition in which the harmony and rhythm remain identical and only the melody is changed by inverting it, apparently does not suffer a change in its expressiveness which distinguishes it in any constant and predictable way from the original.

To some extent these differences in the definiteness of the expressiveness of these elements are the function of the complexity of the element. A change from major to minor is very simple, definite and limited, but there is an infinite variety of ways in which a melodic line may ascend—a long slow ascent, or a succession of short rises, reaching a high point very quickly, a slowly rising pattern of scale-like figures or a series of large skips with repeated climaxes. There is also a variety of ways in which rising and falling of the melody may be combined so as to give prominence to either one quality or the other. In our studies we have chosen 9 compositions which display a variety of melodic patterns, and with these compositions we do not find consistency in the results. More elaborate studies in which certain aspects of rise and fall in melodies are isolated in independence of each other, or studies of some other qualities of the melody, *e.g.* its circular or pendular motion, may demonstrate a much greater importance for the expressiveness of melody.

We believe that refinements in the method of these experiments, especially in the rewriting of the music, will yield results even more striking and clear-cut than those which we are quoting here. Although every possible precaution is taken to eliminate from the experimental material every composition whose two versions differ in more than one respect, it is difficult to anticipate every possible deviation which might affect the results. The two compositions must be matched for tempo, rhythm, harmony, and melody, and the musical beauties of both versions must not be sacrificed in any way.

The results for our studies of rhythm would give much more clear-cut differentiation were it not for one error which occurred in the use of the Prokofieff Gavotte. On every second and fourth beat in the firm version there is a grace-note introducing the chord. This grace-note was not included in the flowing version, and its presence undoubtedly causes the increase in playfulness and gayety which this one composition alone shows for the *firm* version (see Table 1). A further analysis of the studies in harmony would also lead to some further details of expressiveness. A comparison of our complex versions in which the chords were diminished and those in which they were augmented leads us to the conclusion that the augmented dissonance, *e.g.* the nervous modern discords in the modern 'movie' style, makes the music more exciting and impetuous, and that diminished and minor seventh dissonance makes the music more depressed and yearning. A more detailed study of these two factors might establish these differences quite exactly.

There are certainly additional elements of expressiveness which we have not investigated here, and there are other aspects of these four elements which are no doubt of great importance in the expressiveness of music. We have left entirely untouched the effect of tempo, for example, an element which could be investigated by this method with the greatest ease. We have studied only one aspect of rhythm, and left unexplored two aspects which would probably prove to be of importance, the duple *vs.* the triple rhythm, and the smooth *vs.* the jerky rhythms. In our studies of harmony we have still to explore the effect of different kinds of resolutions and different styles of consonance and dissonance, augmented and diminished, etc. The interrelations between these harmonic styles, augmented dissonace in major *vs.* augmented dissonance in minor, may also be explored.

The relative importance of the various elements, tempo, rhythm, harmony, etc., would best be studied by an extension of our method, in which eight or more versions of the same composition would be written, embodying all the possible permutations of several variables. A half-dozen compositions, very carefully selected and treated in this way, would not completely solve all of the problems of expressiveness of music, but would certainly extend our knowledge of the field beyond its present restricted limits.

Summary

(1) In these experiments, we have assumed that there is some systematic symbolism underlying the expressiveness of music, and we have been able to show that the meaning of both short and long compositions, (short themes of 8 to 12 measures and complete musical compositions such as opera preludes and symphonic movements) is generally understood by both trained and untrained listeners. There is striking uniformity and consistency in the interpretations.

(2) The *O*s report their reactions by checking adjectives from a printed list, and the process of combining the adjectives into homogeneous groups for statistical treatment has greatly simplified and clarified the results, and has emphasized the consistency and the systematic character of musical meanings.

(3) We have tried to isolate four of the elements of musical form in

order to discover the meaning or character that is associated with each. The method is that of preparing two versions of a musical composition which differ in one respect only, for example, the rhythm. In this way, the element to be investigated is left in its proper setting, and the stimulus which the *O* hears is real music, not isolated chords, or unsupported melodies.

(4) The elements studied and their associated meanings as determined by this method are as follows: (a) The major mode is happy, merry, graceful, and playful, the minor mode is sad, dreamy, and sentimental, and such qualities as excitement, vigor, dignity, serenity, etc. are not determined by either mode. (b) Firm rhythms are vigorous and dignified; flowing rhythms are happy, graceful, dreamy, and tender, and neither is particularly useful in determining such characteristics as excitement, satisfaction, and serenity. (c) Complex, dissonant harmonies are exciting, agitating, vigorous, and inclined toward sadness; simple consonant harmonies are happy, graceful, serene and lyrical. (d) Differences in expressiveness caused by the rising or falling of the melodic line are not clear-cut, distinct, or consistent. There are tendencies toward the expression of both exhilaration and serenity by the descending melodies, and toward dignity and solemnity by the ascending.

A COMPARISON OF TACTUAL-KINESTHETIC AND VISUAL PERCEPTIONS OF EXTENT AMONG ADULTS, CHILDREN, AND SUBNORMALS

By THEODORA M. ABEL, Sarah Lawrence College

Renshaw found that when children up to the age of puberty were asked to reproduce the location of a stimulated point on the arm,[1] they were more accurate when blindfolded and allowed to indicate with the non-stimulated arm the point on the skin at which they judged the initial stimulus had been applied.[2] Adults, on the other hand, were more accurate when presented visually with an outline drawing of the hand or arm surface (scaled to size), and allowed to indicate the position on the drawing that, in their estimation, corresponded to the point of stimulation on the skin.[3]

These differences in mode of response are not due, according to Renshaw, to the fact that adults are out of practice in making body orientations, as has been suggested by other investigators.[4] Rather, he explains these differences as being dependent on the fact that, in the seeing adult, tactual-kinesthetic localizing movements have suffered a decrement because of the shift to dependency upon vision as the principal or dominant sensory control.[5] Among children there is less interference from vision, so that they function more adequately when relying primarily on tactual-kinesthetic cues.

The present investigation undertakes to ascertain whether this difference in preferred modes of perceiving by children and adults carries over to

* Accepted for publication December 13, 1934.

[1] S. Renshaw, The errors of cutaneous localization and the effect of practice on localizing movement in children and adults, *Ped. Sem.*, 38, 1930, 223-238; S. Renshaw and R. J. Wherry, Studies on cutaneous localization; III. The age of onset of ocular dominance, *ibid.*, 39, 1931, 493-496.

[2] For other studies demonstrating the superiority of children in touch localization see R. E. Dunford, The genetic development of cutaneous localization, *Ped. Sem.*, 37, 1930, 499-513, and his references, 511-512. Also Renshaw's references, *op. cit.*, 237, and footnote 1, 224.

[3] The aid of vision in touch localization among adults had been noted previously by Henri, *Über die Raumwahrnehmungen*, 1898, 129. He found that his *Ss* made smaller errors when indicating on a model where they had been touched rather than when pointing (blindfolded) to the skin surface. Renshaw, however, was the first to approach this aspect of the problem genetically.

[4] J. Peterson, Local signs as orientation tendencies, *Psychol. Rev.*, 33, 1926, 218-236; Dunford, *op. cit.*, 510.

[5] For further evidence for this theory see S. Renshaw, R. J. Wherry and J. C. Newlin, Cutaneous localization in congenitally blind versus seeing children and adults, *Ped. Sem.*, 38, 1930, 239-248.

task situations other than orientation on the skin. Will the visual dominance shown by the adults in making cutaneous localizations manifest itself in other types of perceptual responses and interfere with the smooth functioning of their tactual and kinesthetic mechanisms?

With this object in view, we sought a task having conditions comparable to those set up by Renshaw, in which the subject perceives a stimulus initially through touch, and, secondly, estimates this perception by mechanisms predominantly tactual-kinesthetic in the one situation, and predominantly visual in the other situation.

After sampling various tasks to be performed on both children and adults, the estimation of spatial dimension in its simplest form seemed to us to be a mode of functioning that would meet Renshaw's requirements as well as lend itself to quantitative measurement.

The magnitude of errors made in estimating extents has been a fertile field of investigation in psychophysics. Usually the problem has been one of obtaining the absolute or differential limen in a given sensory field, with different standard lengths. At times, however, interest has been aroused in the question of 'sensory equivalents'; in testing the accuracy with which estimates of extent can be made, using a different sense-modality from the one by which the standard has first been apprehended.

Wundt, Fechner, and Jastrow were among the earlier investigators to concern themselves with this problem.[6] Many discrepancies were found between what Jastrow termed the "receiving sense" and the "expressing sense," and these discrepancies were usually larger when the shift in comparing was made from one sense-modality to another than when the comparisons remained within the same sense-modality. The chief tendency seemed to be that when the change was made from a tactual or a tactual-kinesthetic (active or passive) perception to its visual equivalent, the visual reproduction was generally *underestimated,* so that it measured shorter than the original standard. Fechner found, however, that with longer standards, the visual extents were overestimated.

More recently, Foucault took up this problem.[7] His *Ss* showed more variability of performance: not all made their constant errors in one direction, except when movement was allowed during the tactual perception. In this case, the visual reproductions were those oftener overestimated.

Cutsforth studied the relationship between tactual and visual perceptions, using surface rather than linear estimations.[8] Comparisons between a tactual-kinesthetic perception of area and its visual reproduction revealed gross discrepancies. Both trained and untrained *Ss* tended to underestimate the standards, but less so when

[6] W. Wundt, *Beiträge zur Theorie der Sinneswahrnehmung,* 1862, 34-42; G. T. Fechner, *Elemente der Psychophysik,* 2, 1889, 317-322; J. Jastrow, The perception of space by disparate senses, *Mind,* 11, 1886, 539-554.

[7] M. Foucault, La perception des longueurs par la peau, *Année psychol.,* 28, 1927, 75-97.

[8] T. D. Cutsforth, An analysis of the relationship between tactual and visual perception, *Psychol. Monog.,* 44, 1933 (no. 197), 125-152.

the perceptions were obtained through active movement. We have no indication, however, of what the errors would have been in tactual-kinesthetic reproductions as compared with visual reproductions, although we do know from this experiment that gross errors occur when a visual perception is reproduced visually, as well as when a tactual perception is reproduced visually.

Cutsforth falls into the error of assuming that the magnitude of the objective standard measurement (size of rectangle) is the initial tactual-kinesthetic perception. What he does is to compare the results of functional activity (visual estimate) and an objective standard presented by *E* before any functional activity (tactual-kinesthetic perception) has taken place on the part of *S*. The older investigators frequently fell into a similar error of confusing function with stimulus-pattern, and spoke of an exaggeration or underestimation of visual reproductions of initial tactual patterns, although here again their measurements were of the products of a function as compared with an objective standard.

From the results of the experiments we have cited, it appears that adults are *less accurate* in making *visual* estimates of tactual perceptions of extent than in making *tactual-kinesthetic* estimates of these same tactually apprehended extents. At the start, then, it would seem that in this type of activity the dominant visual sensory control, as suggested by Renshaw, would not play the primary rôle as it did in tactual orienting movements. We do not know, however, to what extent the inaccuracies of visual reproductions will exceed those of tactual-kinesthetic reproductions when conditions are thoroughly controlled, and when the *S*s are allowed to make, in the same experimental hour, both visual and tactual reproductions of initial tactual perceptions.

Furthermore, granted that adults function in a different manner, visually and tactually, how will their performance compare with the performance of children? Will the children do relatively better when they use predominantly tactual and kinesthetic mechanisms, and will the adults do relatively better when employing visual mechanisms as the primary mode of operating? No one as far as we know has approached this problem genetically.

Method. *S* sat facing an upright gray screen 60 by 60 cm. in size, with an aperture in the lower center 23 by 25 cm. This aperture could be closed or opened by means of a black curtain controlled manually by *E*. The stimulus-object to be felt and seen was the edge of a gray cardboard lying in the horizontal plane. Its length was controlled by two parallel stops: one stationary and the other adjustable.

The apparatus was set up in the following manner (Fig. 1). Directly back of the aperture in the upright screen, a wooden box 23 cm. wide, 30 cm. deep, and 6 cm. high, was fastened to the laboratory table. The frame served to support two pieces of gray cardboard. One piece, *A*, with rectangular opening was fixed, the other, *B*, could be moved to the right or left over *A*. The extent of the stimulus-object (*a-b* of the rectangular opening) could be varied from 2.5 to 13.5 cm. The left edge of

the opening *a-c* formed the stationary stop; the right edge of the opening *b-d* formed the adjustable stop. The edge *a-b* measured 2 mm. in thickness and was sandpapered to a smooth finish. It was 9 cm. behind the upright screen.

An extension of *B* to the left (3 cm. wide) beyond the upright screen provided the handle with which *S* adjusted the distance *a-b*. A small block of wood fastened at the rear of the upright screen just above the slide prevented vertical movement. The handle was concealed behind a black curtain fastened at one end to the upright screen.

A millimeter scale was placed on *A* in such a way that the extent of *a-b* could be read from the point at which the edge of *B* fell along the scale.

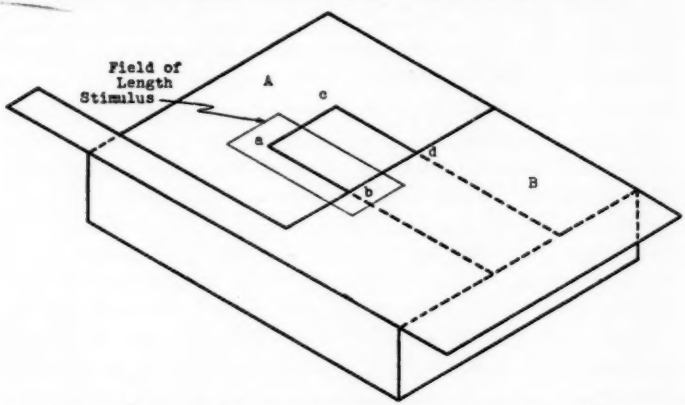


FIG. 1. APPARATUS FOR PRODUCING A STIMULUS-FIELD OF ADJUSTABLE LENGTH
The slide B is adjustable on the fixed cover A, thus providing for variation in length of stimulus a-b. The region outside of the rectangle marked "field of length stimulus" is covered with a black cloth.

Fig. 1 represents the set-up for a left-right (LR) procedure, where *S* begins feeling at *a* and moves his finger along to *b*, thus providing a fixed initial position. For a right-left (RL) procedure, which we also employed, *S* began feeling at *b* and moved his finger along to *a*. Here two other pieces of gray pasteboard were used with *B* stationary and *A* movable. Otherwise, the set-up was the same as that described above.

When actually experimenting, all the surface of *A* and *B* was covered by black cloth exposing only 1 cm. of pasteboard of each stop, *a-c* and *b-d*, and along *a-b*. The far side *c-d* was covered completely, leaving only an aperture of 2.5 cm. The marked rectangle in Fig. 1 indicates the surface exposed. When visual exposures were made, a 3-sided back-drop black curtain allowed for presentation of only the necessary field. The back-drop was low enough for *E* to observe the performance of *S* from above.

Procedure. *S* first extended his right hand and wrist through a slit in the curtain over the aperture in the upright screen, *E* then placed *S*'s finger down at the corner *a*. *S* moved his finger along the edge at an even rate (approximately 10 cm. in 3

sec.) until he reached *b*. He then moved his finger back again from *b* to *a* at the same rate. The positions of the arm and hand were raised enough to allow free movement and to avoid contact with any surface other than the stimulating edge. *S* then lifted his finger and *E* changed the length of *a-b*. The length was varied alternately to 2.5 cm. and to 13.5 cm. In this way adjustments could be made by *S* from smaller to larger in the one case and larger to smaller in the other.

Two procedures were used. (1) In one, where *S* was to reproduce by tactual-kinesthetic means the initial tactual-kinesthetic perception, *E* placed the index finger of *S* back in the initial position *a*. *S* again felt the distance *a-b*. At the same time he extended his left hand and wrist through a slit with flap in the curtain at the left, and adjusted the length of *a-b* by moving the slide to the left or right, *S* was allowed to feel as long as he wanted (up to 15 sec.), until the distance *a-b* felt as long as it had felt in the first presentation. The distance was then read on the scale by *E*. For the right-left procedure the index finger was initially placed at *b*. Only one procedure was used with an individual *S*. (2) In the other procedure, where *S* was to reproduce by visual means the initial tactual-kinesthetic perception, *S* removed his arm from behind the curtain, which was then raised, exposing *a-b*. *S* then adjusted the length *a-b* until the distance from *a* to *b* appeared visually to be equal to his tactual experience. Again *S* was given as much time as he desired (up to 15 sec.) for making his estimate. When the adjustment had been made, the curtain was lowered and the distance was read off the scale.

The first procedure will be designated as TT, and the second as TV. The stimulus-distances used were 5, 7, 9, and 11 cm. Each *S* was given 30 trials—6 preliminary and 24 in the main series—by each of the two procedures.[8] In the main series the TT and TV procedures were alternated after every 4 trials, each of the four stimuli, 5, 7, 9, and 11 cm. being presented once, in a chance order, for each TT and TV presentation. In this way a total of 12 trials for each procedure, and 3 trials for each length was obtained. The initial mode of presentation was alternated with each *S*, so that one began the main series with four trials by the TT method followed by 4 trials of the TV method, and the next *S*, conversely, began the series with the TV method followed by the TT method. The left-right and the right-left procedures were alternated with each successive *S*.

S was told at the beginning, "This is an experiment or game to see how well you can tell how long lines are when you feel them and when you see them." He was then given the specific instructions for the two methods TT and TV. The preliminary trials were not started until it was certain that *S* understood the nature of the task. During the experiment the instructions for each specific step were repeated before each trial. *S* was not told anything about the accuracy of his performance, although for encouragement *E* frequently made the comment 'Good' or 'Fine.' All unsolicited comments on the *S*s about the ways in which their activity was carried on were recorded by *E*.

Each *S* was tested individually, the experimental period lasting 30 to 40 min. The experimental room was illuminated by daylight, controlled only to the extent of insuring good visibility. No left-handed *S* or none with marked visual defects was used. The author of this paper acted as *E* throughout the experiment.

---

[8] In preliminary experiments we ascertained that no significant practice effect occurred by either the TT or TV method up to 25 trials for each.

Experiment I. Normal Children and Adults

The adult *Ss* were 50 undergraduate women, median age 19.4 yr., range 16 to 21 yr. Fifty children (27 boys and 23 girls) also served as *Ss*, 15 between 9 and 10 yr. old, 21 between 10 and 11 yr. old and 14 between 11 and 12 yr. old.[10] The intelligence level of all the children was normal or superior as measured by various standard tests. To 34 of these *Ss* had been given the Terman-Binet tests. Among these 34 the IQ scatter was from 0.93 to 1.56, with a median of 1.14. The experiment was carried on during the winter and spring terms of 1933-34 at Sarah Lawrence College, the Bronxville Public School, and the Friends' Seminary, New York City.[11]

Results. Scoring was made by computing the errors between the objective measurements of the reproduction and the standard initially presented, averaging these errors, without regard to signs, for each of the 4 standard lengths used, and converting this average error into percentages. Table I shows these scores. Scores for standard 11 cm. by the TV method do not adequately represent the reactions of the *Ss*, as a constant error occurred here: the *Ss*, particularly the adults, frequently tried to make visual adjustments beyond the limit (13.5 cm.) set by the apparatus.

TABLE I

Percentages of Errors Made by 50 Children and 50 Adults
(3 trials made by every *S*)

Method	Standard cm.	Children		Adults	
		Av. score	P.E.	Av. score	P.E.
TT	5	24.96	6.84	13.58	4.73
TV	5	36.16	15.08	23.68	11.17
TT	7	17.2	5.71	14.78	4.73
TV	7	24.16	8.52	22.4	8.91
TT	9	13.84	4.55	12.08	3.69
TV	9	18.72	6.86	17.44	7.43
TT	11	17.2	3.85	11.84	3.6
TV	11	13.6	4.71	12.8	4.48

The adults were less influenced by difference in magnitude than the children. By the TT method, their performance was quite similar on the three sizes, 5, 7, and 9 cm., and their tendency to improve by the TV method as the magnitude increased was only slight. The children, on the other hand, had more difficulty with the shortest length by both the TT

[10] The experiment was tried on a few 8-year-old children, but the directions seemed too difficult for them.

[11] Thanks are due Mr. Beatty, Principal of the Bronxville Public School, and Mr. Messner, Principal of Friends' Seminary, New York, for enabling us to conduct the experiment in their schools. Thirty children from the Bronxville school and 20 from Friends' Seminary were given the experiment.

and the TV methods than with the longer lengths. Jastrow and Foucault both found that the errors decreased with increase in magnitude of the standards.[12]

TABLE II

PERCENTAGES OF ERRORS MADE BY CHILDREN AND ADULTS TOGETHER WITH THE MEASURES OF RELIABILITY OF THE DIFFERENCES

Ss	No. of scores	Method	Standards cm.	Av. score	P.E.	Diff.	P.E.$_{diff.}$	Crit. R.
Children	150	TT	5, 7, 9	19.33	6.38	1.90	.63	9.2
Adults	150	TT	5, 7, 9	13.49	4.48			
Children	150	TV	5, 7, 9	26.34	11.84	2.36	1.25	4.1
Adults	150	TV	5, 7, 9	21.18	9.48			
Children	100	TT	7, 9	18.42	6.79	2.72	.79	6.6
Adults	100	TT	7, 9	13.19	4.07			
Children	100	TV	7, 9	21.44	7.98	−0.46	1.16	1.3
Adults	100	TV	7, 9	19.92	8.44			

The adults did better than the children by both methods. Table II presents the average scores for the two groups when the scores, first on three standards and then on two standards, are thrown into one distribution.[13] The second comparison, using only standards 7 and 9 cm., was made because of the poor and variable performance of the children with standard 5 cm. As can be seen from the critical ratios, the difference between the performance of the children and the adults, when the results of the three standards were computed, are statistically reliable (100% certainty by the TT method and over 200% certainty by the TV method). When the results for standards 7 and 9 cm. were computed, the adults still performed reliably better than did the children by the TT method, but by the TV method they did not do so (only 89% chances for a true difference between the averages). Both groups were less variable in the TT performance than in the TV performance. The children were 83% and the adults 85% as variable in the former as they were in the latter.[14]

We compared the TT performance of the individuals who initially moved their index finger from left to right and those who initially moved it from right to left. (Table III). The difference between the LR and the RL scores for the children was not significant. For the adults the LR procedure was somewhat easier than the RL procedure, although the difference between the averages does not meet the requirements of complete statistical

[12] Jastrow, *op. cit.*, 544; Foucault, *op. cit.*, 86.

[13] As the scores on standard 11 cm. by the TV method had to be omitted, because of the constant error, we excluded in our comparisons also scores on standard 11 cm. by the TT method.

[14] The formula employed in obtaining these percentages was: $V = 100\ S.D./av.$

reliability (a critical ratio of 2.4 giving a 95% chance for the difference to be a true one).

TABLE III

PERCENTAGE OF ERRORS FOR LEFT-RIGHT (LR) AND RIGHT-LEFT (RL) PROCEDURES, AND MEASURES OF RELIABILITY OF DIFFERENCES

Ss	Procedure	No. Scores	Av. score	P.E.	Diff.	P.E.$_{diff.}$	Crit. R.
Children	LR	100	17.88	5.69	0.40	.83	.95
Children	RL	100	18.68	6.09			
Adults	LR	100	12.21	3.62	0.78	.57	2.4
Adults	RL	100	13.55	4.40			

In both the LR and RL procedures the *Ss* felt not only in the one direction but also in the opposite direction, that is, after crossing the extent once with the index finger they were allowed to return to the point where they had started. Consequently, in the RL procedure cues could be obtained from the preferred direction of movement among right-handed *Ss*. Possibly larger differences in scores of the LR and RL procedures might have occurred if each trial had been limited to movement in one direction.

As indicated under 'Procedure,' one half of the *Ss* began the main experimenting with 4 trials by the TT method, followed by 4 trials by the TV method; while the other half had the initial order of presentation reversed. The scores for the TT_1 and TT_2 methods (TT presented first or second) and for the TV_1 and TV_2 methods (TV presented first or second) were computed separately (see Table IV).

TABLE IV

PERCENTAGE OF ERRORS FOR ORDER OF TT AND TV PRESENTATIONS, AND MEASURES OF RELIABILITY OF DIFFERENCES

Ss	Method	Av. score	P.E.	Diff.	P.E.$_{diff.}$	Crit. R.
Children	TT_1	20.54	6.66	2.63	1.02	2.6
	TT_2	17.91	5.89			
	TV_1	26.0	10.48	0.8	1.76	0.45
	TV_2	25.2	11.15			
Adults	TT_1	14.97	4.16	2.00	0.65	3.1
	TT_2	12.97	3.75			
	TV_1	19.35	8.26	−2.92	1.45	2.0
	TV_2	22.27	9.41			

For both children and adults the TT method was somewhat easier when it was initially presented.[18] The differences between the averages are not, however, reliable statistically (95% chances in the case of the children and 98% chances in the case of the adults for true differences), but the trend of behavior, however, is indicated. Now, proceeding by the tactual-kinesthetic method was undoubtedly a more direct and simple task on the average for the adults and to a certain extent for the children, as we shall see further in our discussion of the relative ease with which each type of performance was carried out (see also Table II). Hence the

[18] Three practice trials for each of the two methods were given each *S* prior to this presentation, so that 'initial performance' here does not mean the first attempt.

contrast of harder (TV) to easier task (TT) might possibly have affected the manner in which the *Ss* reacted, or at least have influenced their self-instructions in some way. A number of unsolicited comments from the adult *Ss* give a hint as to some of the ways in which the tasks were met. The children make no comments of any value here.

The visual performance was not so much affected by the order of its appearance. At least for the children, the difference between TV_1 and TV_2 scores was not statistically reliable (only 62% chances for a true difference), while the difference for the adults was somewhat more significant (91%). The adults on the whole were more influenced than were the children by the order of presentation of the two methods, TT and TV. Sophisticated *Ss* like the adults would naturally give themselves more formal self-instruction in carrying out a required task of some difficulty, and consequently their performance might have been more differentially influenced than that of naïve *Ss*. This factor might have accounted also for the differential performance of the adults on the LR and RL procedures.

In psychophysical work, one type of space error is taken into account by having the *S* make adjustments both in the direction of longer to shorter and in that of shorter to longer. This we did, as stated in the procedure, by alternating the adjustments 'longer to shorter,' (LS) and 'shorter to longer,' (SL) in each successive trial.[16]

In both the LS and SL procedures by the TT method, where the length presented for adjustment was 13.5, *S* could move his index finger from the starting point across a desired distance without encountering any obstacle, so that he was able the more directly to reproduce the initial tactual-kinesthetic perception of extent. In the SL procedure by the same method, where the length presented for adjustment was 2.5 cm., *S* had to react differently, for he encountered the stop (*b-d* or *a-c*, Fig. 1) before he could make the full extent of movement desired.

We computed the percentage errors for the reproductions by the two procedures, LS and SL, and averaged the results for the two shorter lengths, 5 and 7 cm., and the two greater lengths, 9 and 11 cm., separately (Table V).[17]

The results indicate that, for the two greater lengths, when both children and adults were required to make the adjustment SL, their errors were larger on the average than when they had to make the adjustment LS. For the two shorter lengths, the children were similarly influenced, incurring larger errors when they had to make the greater adjustment LS and smaller errors when they had to make the smaller adjustment SL. The differences between the averages in these three comparisons are completely reliable. The reactions of the adults on the two shorter lengths, however,

[16] Cutsforth eliminated the influence of size entirely by reducing to zero the visual diaphragm to be adjusted, thus having no surface exposed prior to the judgment. We were interested in comparing, however, the differences in the perceptions of increasing and decreasing extents.

[17] Standard 11 cm. was included here, as the results were not to be compared directly with those by the TV method. In this case the percentages of error were obtained from each individual trial instead of from the average of three trials, as in the other estimates.

were the same in both the LS and SL procedures, regardless of the distance through which the adjustment had to be made. Apparently, they were able to make finer distinctions here and compensated for the direction and extent of adjustment to be made.

Judging from the results as a whole, the *S*s were more influenced by the extent of the perception (very long or very short) when setting out to make the adjustment, and by the amount of movement required of the left arm, rather than by the fact that in the one case, LS, the index finger had a free sweep of movement, while in the other case, SL, the finger was blocked by the stop before it could move to the full distance desired.

As the performance with standard 11 cm. could not be included by the TV method, we compared the scores on the LS and SL procedures for a short (5 cm.) and a long (9 cm.) standard (Table V). The reactions of

TABLE V

PERCENTAGE OF ERRORS FOR LONGER-SHORTER (LS) AND SHORTER-LONGER (SL) ADJUSTMENTS, AND MEASURES OF RELIABILITY OF DIFFERENCES

Ss	Method	Standards cm.	Adjustment	No. of scores	Av. score	P.E.	Diff.	$P.E._{diff.}$	Crit. R.
Children	TT	5, 7	LS	155	23.57	11.0	6.28	1.1	5.6
		5, 7	SL	145	17.29	8.64			
		9, 11	LS	145	12.5	6.18	−8.77	0.84	10.4
		9, 11	SL	155	21.27	8.33			
Adults	TT	5, 7	LS	155	14.07	7.43	−.07	0.77	0.09
		5, 7	SL	145	14.14	5.88			
		9, 11	LS	145	9.66	5.33	−5.89	0.65	8.4
		9, 11	SL	155	15.15	6.09			
Children	TV	5	LS	98	39.9	17.95	11.35	2.6	4.4
		5	SL	52	28.55	13.48			
		9	LS	76	16.01	7.56	−4.43	1.3	3.4
		9	SL	74	20.44	8.68			
Adults	TV	5	LS	98	26.02	13.62	6.51	2.2	2.9
		5	SL	52	19.51	11.87			
		9	LS	76	19.16	9.39	−0.99	1.48	0.67
		9	SL	74	18.17	8.82			

the children were similar to their reactions by the TT method: the larger percentage errors were made, the greater the amount of adjustment required (differences in averages between LS and SL completely reliable on standard 5 cm. and practically so on standard 9 cm). The adults were very much less influenced by the direction of adjustment (differences in averages between LS and SL are not reliable). As was to be expected, both groups were more affected by the LS procedure on standard 5 cm., where the adjustment was 8.5 cm., than by the SL procedure on standard 9 cm. where the adjustment was only 6.5 cm.

The question of whether a sex difference in performance was shown by the children in this experiment was considered, especially in view of the fact that boys are likely to have had more manual training than girls. In comparing the scores made by the two groups of children, we found that the boys tended to do better than the girls by the TT method (Table VI). The difference, however,

TABLE VI

PERCENTAGE OF ERRORS FOR BOYS (27) AND GIRLS (23), AND MEASURES OF RELIABILITY OF DIFFERENCES

Ss	Method	No. scores	Av. score	P.E.	Diff.	P.E.$_{diff.}$	Crit. R.
Boys	TT	81	18.12	5.75	−2.52	1.05	2.4
Girls	TT	69	20.64	6.9			
Boys	TV	81	24.64	11.2	−3.01	1.85	1.8
Girls	TV	69	27.65	11.35			

does not reach statistical reliability, the critical ratio being only 2.4 (95% chances for a true difference). With a larger number of cases, this difference might become more significant. By the TV method, the performance of the boys and girls was much more comparable. The difference in the average scores was not reliable (only 89% chances for a true difference).

Correlations between intelligence test scores and performance by the TT and the TV methods were obtained by the method of rank differences for 35 adults who had been given the Thorndike Scholastic Aptitude Examination and for 32 children who had had the Stanford-Binet Test.

The Pearson *rs* for these correlations are:

Adults	Thorndike score and TT score $r = -.14$
	Thorndike score and TV score $r = +.13$
Children	IQ and TT score $r = +.01$
	IQ and TV score $r = +.21$

These correlations are negligible. The better performance of the adults when employing tactual-kinesthetic mechanisms, therefore, was not necessarily dependent on their superiority of comprehension and thought as measured by an intelligence test. The task of making visual reproductions, in spite of their "better tools for thinking," was apparently harder for the adults, so that their performance did not greatly surpass that of the children, except on the shortest length.

We next obtained the percentage frequencies with which the measurements of the TT and the TV reproductions exceeded (>), fell short of (<), or equalled (=) the four standard lengths. Thus:

Ss	Method	>	<	=
Children	TT	.35	.62	.04
Adults	TT	.31	.62	.07
Children	TV	.58	.38	.02
Adults	TV	.67	.31	.02

The results for the two modes of reproduction are the reverse of each other for both adults and children. For the TT method the ratio of the

frequency of measurements of the reproductions that fall short of, to the frequency of those that exceed, the standard lengths is approximately 2:1. For the TV method this ratio is reversed and becomes 1.2 for adults and 2:3 for children. In other words, by the TV method the measurements of the reproductions oftener exceed than fall below the standard lengths when the TV method is used, and oftener with the adults than with the children.

When the number of cases was computed separately for the different standards, however, three exceptions to the constant errors indicated above occurred (Table VII; see rows opposite 'mean'). (1) With the standard 5 cm. the children reacted by the TT method as they did by the TV method: the percentage greater than the standard exceeded those less than the standard by the ratio of 7:4. (2) For the adults by the TT method, the percentage of cases exceeding or falling short of the 5-cm. standard were approximately equal. (3) For the children by the TV method, the percentage exceeding or falling short of the 11-cm. standard were also approximately equal.

Underestimations of the reproductions by the TT method have been reported by other investigators. Kramer and Moskiewicz had their *Ss* make arm movements across a given distance on a table, with range of movement 40 to 160 mm., and noted a persistent tendency for the arm to move a shorter distance in the reproduction than in the initial movement.[18] Jastrow on the contrary found that his *Ss* exaggerated the reproduced movement.[19] Budkiewicz had *Ss* make several reproductions of a 6 mm. standard in the visual, tactual, and kinesthetic sense-modalities.[20] In the kinesthetic modality, where the performance was the more comparable to ours, her *Ss* also greatly exaggerated the movement of the reproduction. A study by Calabresi on the comparative tactual-kinesthetic perception of magnitude and form among children and adults is of interest here, although the task required was quite different from ours: manipulation of objects.[21] He noted that among both normal and abnormal children the tendency was to judge the size of an object as less than it actually measured. Adults showed the same tendency but to a less degree.

The results for our TV method, on the other hand, where adjustments longer than the standards occurred oftener than adjustments shorter than the standards, are similar to the results of Foucault and those of Fechner for the greater lengths, but contrary to the results of the other investigators referred to earlier in this paper.[22] In none of these experiments has the nature of the task or the conditions under which it was performed been very similar. It is not surprising, therefore, to find contradictions in results.[23]

---

[18] F. Kramer and G. Moskiewicz, Beiträge zur Lehre von den Lage- und Bewegungsempfindungen, *Zsch. f. Psychol.*, 25, 1901, 114-115.

[19] Jastrow, *op. cit.*, 544.

[20] J. Budkiewicz, Étude expérimentale sur les processus de mesure spatiale linéaire, *Année psychol.*, 27, 1926, 94-173.

[21] R. Calabresi, La percezione tattilo-cinetica delle grandezze e delle forme negli adulti e nei fanciulli, *Riv. sci. appl. all' educ. fis.*, 2, 1931, 1-37.

[22] Foucault, *op. cit.*, 86; Fechner, *op. cit.*, 320 f.; Wundt, *op. cit.*, 35; Jastrow, *op. cit.*, 546; and Cutsforth, *op. cit.*, 135.

[23] Wundt and Fechner had their *Ss* make estimates of the extent of passive tactual

We next found that the number of cases in which these measurements for the reproduced perceptions exceeded, fell short of, or equalled the initial standard lengths was influenced by whether the adjustments were made from longer to shorter or from shorter to longer, although this influence varied with the different standards used, the two modes of procedure, TT and TV, and the different groups of *Ss* (Table VII). In the LS adjustments by the TT method, the measurements of the reproductions oftener exceeded than fell short of the standards for the adults on lengths 5, 7, and 9 cm. and for the children only on lengths 5 and 7 cm. The reverse results were obtained for the adults on length 11 cm. and for the children on both the longer lengths, 9 and 11 cm. In the SL adjustments by the TT method children and adults reacted similarly. The results indicate that the frequencies of measurements overestimating the standards greatly outnumbered the frequencies underestimating them. This held true for all 4 standards.

Turning to the TV method, we find that for both children and adults, when they adjusted from longer to shorter, LS, the frequencies of measurement overestimating the standards greatly exceeded those of underestimation for all standards.

In adjusting by the SL procedure, the adults again reacted as by the LS procedure for all lengths, although here the frequencies of overestimation were not as excessive as in the former LS procedure. The children in the SL procedure, on the other hand, reacted as in the LS procedure, only on the shortest length. For the two middle lengths, 7 and 9 cm., the frequencies of over- and underestimations of the standards were about equal, while for the 11 cm. length the frequencies of overestimations exceeded those of underestimations in the proportion of 3:2. In general results for children and adults were fairly similar, with one quite marked exception. For the adults in the TV method, by the SL as well as by the LS procedure, the frequencies of overestimation adjustments, greatly exceeded those of underestimation.

The other problem investigated in this experiment was: granted that adults function more adequately than children by the TT and TV methods, what will the nature of their relative performance be? What will be the discrepancies in the results of the two methods for the adults as compared to the children?

---

pressures. Jastrow required, in one procedure, active movement of the arm. His *Ss* made visual estimates of these movements by selecting lines of given length. Foucault had his *Ss* make visual estimates by delimiting a distance on a line traced in advance. Cutsforth worked with surfaces rather than extents.

TABLE VII

PERCENTAGE OF ESTIMATIONS GREATER THAN (>), LESS THAN (<), AND EQUAL TO (=) STANDARDS FOR LONGER-SHORTER (LS) AND SHORTER-LONGER (SL) ADJUSTMENTS FOR CHILDREN, ADULTS, AND SUBNORMALS FOR TT AND TV METHODS

Ss	Method	Adjust.	5 cm.			7 cm.			9 cm.			11 cm.			Total 4 standards		
			>	<	=	>	<	=	>	<	=	>	<	=	>	<	=
Children	TT	LS	.80	.20	0	.68	.28	.04	.43	.52	.05	.31	.62	.07	.57	.39	.04
		SL	.29	.67	.04	.25	.71	.04	.09	.71	.04	0	.99	.01	.15	.82	.03
		Mean	.54	.43	.02	.46	.49	.04	.26	.61	.04	.15	.80	.04	.36	.60	.03
	TV	LS	.76	.20	.04	.84	.11	.05	.71	.27	.02	.53	.40	.07	.71	.25	.04
		SL	.69	.25	.06	.49	.49	.02	.47	.49	.04	.38	.60	.02	.49	.48	.03
		Mean	.72	.22	.05	.66	.30	.03	.59	.38	.03	.45	.50	.04	.60	.36	.03
Adults	TT	LS	.56	.26	.18	.60	.33	.07	.51	.45	.04	.41	.55	.04	.52	.39	.09
		SL	.25	.67	.08	.10	.86	.04	.12	.86	.02	.06	.93	.01	.12	.84	.04
		Mean	.53	.40	.13	.35	.56	.05	.31	.65	.03	.23	.74	.02	.32	.61	.06
	TV	LS	.55	.39	.06	.76	.20	.04	.74	.26	0	.74	.24	.02	.70	.26	.04
		SL	.54	.40	.06	.62	.37	.01	.59	.37	.04	.57	.43	0	.59	.39	.02
		Mean	.54	.39	.06	.69	.28	.02	.66	.31	.02	.65	.33	.01	.64	.32	.03
Sub-normals	TT	LS	.69	.28	.03	.63	.33	.04	.46	.51	.03	.38	.59	.03	.55	.42	.03
		SL	.43	.49	.08	.23	.73	.04	.12	.85	.03	.09	.91	0	.20	.77	.03
		Mean	.56	.38	.05	.43	.53	.04	.29	.68	.03	.23	.75	0	.37	.59	.03
	TV	LS	.80	.20	0	.86	.14	0	.82	.18	0	.71	.29	0	.79	.21	0
		SL	.85	.15	0	.60	.35	.05	.61	.38	.01	.63	.36	.01	.63	.32	.05
		Mean	.82	.17	0	.73	.24	.02	.71	.28	.01	.67	.32	.01	.71	.26	.02

We computed the reliability measures of the actual differences between the scores by the TT and the TV methods for both children and adults.[24] The results were as follows:

Ss	Standards (cm.)	TT Av. score	TV Av. score	Diff.	P.E.$_{diff.}$	Crit. R.
Children	5, 7, 9	19.33	26.34	7.01	1.1	6.3
Adults	5, 7, 9	13.49	21.18	7.69	.7	11.0
Children	7, 9	18.42	21.44	3.02	1.05	2.9
Adults	7, 9	13.19	19.92	6.73	.94	7.1

In the first comparison, the differences in averages are completely reliable for both groups. The critical ratio for the adults, however, is almost twice that for the children. In the second comparison, the reliability of the difference remains highly significant for the adults, whereas for the children it has become very much less so, the critical ratio being 2.9 instead of the 4 necessary for complete reliability.

We obtained the same results for the children when standard 11 cm. was included with the standards 7 and 9 cm. The difference between the TT and TV scores was again not completely reliable statistically (TT av. 16.7 ± 4.91, TV av. 18.72 ± 7.55, P.E.$_{diff.}$ 0.73, crit. ratio 2.8). This comparison could not be made with the scores of the adults, as we have already indicated, because of the constant error for them on standard 11 cm. by the TV method.

TABLE VIII

PERCENTAGE OF ERRORS FOR DIFFERENT AGE-GROUPS

Age (yr.)	No. of scores	TT method		TV method	
		Av. score	A.D.	Av. score	A.D.
9	45	19.47	5.55	22.13	6.24
10	63	19.56	3.21	27.04	12.03
11	42	18.59	3.61	28.62	8.9

In order to find out whether the factor of age entered into the performance of the children, we computed the scores separately for the three age levels, 9, 10, and 11 yr. (Table VIII). The three groups did about equally well by the TT method, but the youngest group tended to do better than the two older groups by the TV method. Obviously the number of cases in each age-group is too small for statistical reliability. We feel it is of interest, however, to note the trend in the performance of the older children when they employed predominantly visual mechanisms. The 10- and 11-yr.-old children tended to behave more like the adults and to show a greater discrepancy between their performance by the TT and the TV methods than did the 9-yr.-olds.

[24] Product-moment *rs* between the TT and TV scores were + 0.16 ± .09 for the children and + 0.35 ± .09 for the adults. As these *rs* are so low, we considered their influence negligible and employed the formula (assuming zero *r*) P.E.$_{diff.}$ = [(P.E. av. 1)^2 + (P.E. av. 2)^2]$^{1/2}$ for obtaining reliability measures of the differences between the average scores.

Let us consider some of the reasons that may account for the differences in behavior of the children and adults. In the first place, the task in the TT method required fine discriminations of small arm-movements. The numerous manipulative functions of the hand and fingers with accompanying arm-movements would help in distinguishing small differences in tactual-kinesthetic perceptions of extent. The adults, therefore, would have an advantage in this type of activity. If, as we know, adults are more practiced than children in eye-hand coördinations, we should also expect them to make visual reproductions of tactual-kinesthetic perceptions a good deal better than children can. They did not do so, as we have seen. Two reasons may account for this relatively poor performance by the TV method on the part of the adults.

One is that in the task required in this experiment, the temporal order of the functioning of the 'eye-hand' mechanisms was different from the more usual sequence met with in daily life. Among seeing individuals the 'eye' guides the 'hand' more habitually than the 'hand' the 'eye.' Our visual space perceptions either precede or accompany our tactual-kinesthetic perceptions and give cues that make clearer and more explicit the tactual and kinesthetic space relationships in our phenomenal world.[25] By the TV method, visual representations of tactual-kinesthetic space perceptions had to be construed from cues coming not only from the less practiced source but from that furnishing less finely differentiated perceptions. It follows that the adults had not the same advantage in the TV method as they had in the TT method of functioning.

Training in specific fields would influence the manner in which the visual reproductions were made. Mechanical drawing, for one thing, gives the type of practice necessary for making visual reproductions more comparable to tactual-kinesthetic reproductions.[26]

Spontaneous comments from many of our adult *Ss* revealed the fact that the task of shifting from touch to vision was the more difficult and one in which they were not practiced. Among the children no spontaneous comments were elicited that suggested any difficulty in the visual task. The task was either not so hard for them or they were not hampered in their performance by being aware of the nature of the difficulty. In any case, the children met the situation more easily than did the adults, judging from their behavior during the experimental hour. Furthermore, the superior performance of the 9-yr.-old children by the TV method (Table VIII) suggests that the younger the child, the easier is the shift from a tactual-kinesthetic perception of extent to its visual reproduction.

In addition to lack of training or practice and consequent insufficiency of adequate cues in shifting from tactual-kinesthetic to the visual estimates of extent, the chief factor that accounted for the type of performance made by the *Ss* was the manner in which the initial tactual-kinesthetic perception

[25] For an experiment in which visual cues aided in the building up of tactual space perceptions, see T. M. Abel, The influence of visual and auditory patterns on tactual recognition, this JOURNAL, 46, 1934, 443-447.

[26] We excluded from our experiment *Ss* who had had any extended training in mechanical drawing, as we found that two *Ss* in our trial series with training in this field found it very easy to shift from tactual-kinesthetic perceptions to visual reproductions.

aroused in the individuals a tendency to act in a specific way when making the visual reproductions. As we have seen, the greater number of measurements of the visual reproductions exceeded the standards in length. This mode of activity occurred among both children and adults, but more predominantly among the adults for all 4 standard lengths and in both modes of adjustments, from longer to shorter, LS, and from shorter to longer, SL (Table VII).

Many comments of the adults indicated that they were aware of a discrepancy between the extent of their tactual-kinesthetic perception and their visual reproductions. That visual cues play a rôle in tactual-kinesthetic perceptions for seeing individuals goes without saying, since complete separation of function in different active sense-modalities would be a fiction in an integrated organism. These visual aspects of tactual perception have been observed and recorded by various investigators both in studies on tactual localizations on the skin and in perceptual estimates of extents and surfaces."[27] In fact, according to the Gestalt point of view, all tactual and kinesthetic space perceptions are localized in experienced space more or less in relation to visual experiences.[28] As we have seen, many of our adult *Ss* stated that they had "ideas" about how the "distance would look," but these visual cues were not always correct or adequate.[29]

In some preliminary experiments we allowed the *Ss* to use vision when they were making tactual-kinesthetic reproductions. Their comments here were that looking at the stimulus-object hindered rather than helped their performance, for "it looked smaller than it felt." What they did was to shut their eyes or look away and make their adjustments through touch and kinesthesis. We had no information from the children as to the ways in which they anticipated their visual perceptions. At all events we can assume that they were more naïve than the adults and consequently less aware of discrepancies between the way extents 'feel' and 'look.' If the more natural mode of functioning among *Ss* of any age, however, is for visual reproductions of tactual-kinesthetic perceptions of extent to measure longer than standards, then children should show the same tendency in a more exaggerated manner than do adults, who were frequently 'aware' of a discrepancy between touch and vision and undoubtedly tried to compensate by making visual reproductions appear shorter than they really considered they should 'look.' That such was not the case we have seen from the results of this experiment. Children did show the same tendency as did the adults, but not to the same extent, except when reacting to the shortest standard, 5 cm.

In Experiment II we shall consider in greater detail some of the reasons for the discrepancy between touch and vision as the performance of subnormal *Ss* brought out certain aspects of this problem more significantly.

---

[27] Cutsforth, *op. cit.*, 147-48. For further references on experiments on cutaneous localization and on the perception of spatial relations, in which the dependence of tactual on visual perceptions or on visual imagery is shown, see Cutsforth, *op. cit.*, 151-52; Wundt, *op. cit.*, 39, and Fechner, *op. cit.*, 318.

[28] W. Köhler, *Gestalt Psychology*, 1928, 231.

[29] Henri, *op. cit.*, 120-22, mentioned the fact that visual imagery interfered with the visual perceptions of the model of the arm.

Experiment II. Subnormals

As we have seen previously, scores on the tactual-kinesthetic reproductions and on the visual reproductions did not correlate in any significant manner with scores on intelligence tests. In view of this fact, we considered the problem of the way in which a group of adults would react who had remained fairly immature in the development of the more difficult modes of functioning, comprehending, and thinking ('intelligence'). Would a subnormal group of this kind behave like normal children, showing less differences between their tactual-kinesthetic and their visual reproductions, or would they react more like normal adults revealing greater discrepancies in performance between the one type of reproduction and the other?

If practice and training in manipulative activities influence the performance, as we have suggested, then a subnormal group should make fewer errors than children by the TT method, but react in a manner more or less comparable to both children and normal adults by the TV method.

To investigate this problem, we decided to repeat our experiment on a group of 50 girls, who had reached maturity (at least 16-yr. old), but whose mental ages as measured on a standard intelligence test were 9-, 10-, or 11-yr. old. Thus the 'subnormal' group would approximate in 'mental age' the chronological age-levels of our group of 50 normal children.

We selected our *Ss* in the following manner. We first gave the Otis Self-Administering Examination, Intermediate Form B, to all the pupils (220) in the Trade Extension Classes of the Manhattan Industrial High School for Girls. No one of these girls had completed successfully the 8B grade in the public schools, the majority of them having come from the ungraded or adjustment classes. In the Extension Classes the girls were being taught the simpler processes of dressmaking, millinery, novelty work, and machine operating. They were not considered sufficiently 'intelligent' to carry on the regular work in the skilled trades of the high school. From among these 220 girls we chose 50 whose chronological ages ranged from 16 yr. 0 mo. to 17 yr. 11 mo., and whose mental age ratings on the intelligence test ranged from 9 yr. 0 mo. to 11 yr. 11 mo. Of the 50 *Ss*, 17 had mental age ratings between 9 and 10 yr., 17 between 10 and 11 yr., and 16 between 11 and 12 yr.[20]

The experiment was carried on during the spring term of 1934 at the Manhattan Industrial High School.[21] The author acted as *E* throughout. The procedure and method of computing results were like those described in Experiment I. We computed the error in percentage for the 3 lengths, 5, 7, and 9 cm., and again for the 2 lengths, 7 and 9 cm., with the following results:[22]

[20] No older girls were available, as they are not allowed to remain in the school after the age of 18 yr.

[21] Thanks are due Miss Florence Marshall, Principal of the Manhattan Industrial High School for Girls, and Miss Mary Collins, Director of the Extension Classes, for making it possible for us to carry out this experiment.

[22] The constant error with standard 11 cm. by the TV method was apparent also in the performance of the subnormal group. Consequently, the scores on this standard were excluded here.

Method	Standards (cm.)	No. of scores	Av. score	P.E.	P.E.$_{diff.}$	Crit. R.
TT	5, 7, 9	150	17.24	6.13	1.12	10.7
TV	5, 7, 9	150	29.18	12.24		
TT	7, 9	150	15.65	4.30	1.06	9.3
TV	7, 9	150	25.53	9.64		

The subnormal girls did relatively much better by the TT method than by the TV method. The critical ratios greatly exceeded the requirements of complete statistical reliability.[33] In this respect, the subnormal group behaved in a manner more comparable to that of the normal adults than that of the children, especially in the second comparisons (Standards, 7 and 9 cm.), showing, as did the college students, greater discrepancy in performance between the tactual-kinesthetic and the visual reproductions than did the children (see crit. ratios Table II, p. 275).

We next compared the results among the three groups of the TT and TV methods respectively. The scores by the TT method of the subnormals hold an intermediate position between those of the children and the college students (Table IX). The differences between the averages are com-

TABLE IX

Difference between Average Scores of Subnormals, and of Children and Adults and Measures of Reliability of Those Differences

Compar. of subnormals with	Standards (cm.)	Method	Diff.	P.E.$_{diff.}$	Crit. R.
Children	5, 7, 9	TT	+2.09	.72	2.9
Adults	5, 7, 9	TT	−3.74	.61	6.1
Children	5, 7, 9	TV	−2.83	1.38	2.0
Adults	5, 7, 9	TV	−8.0	1.26	5.3
Children	7, 9	TT	+2.77	.81	3.4
Adults	7, 9	TT	−2.45	.59	4.2
Children	7, 9	TV	−4.08	1.25	3.3
Adults	7, 9	TV	−5.6	1.28	4.4

pletely reliable for the subnormals and adults and almost so for the subnormals and children (97% chances for lengths 5, 7, and 9 cm., and 99% chances for lengths 7 and 9 cm., for true differences).

By the TV method, on the other hand, the scores of the subnormal group are larger than both those of the adults (differences 100% reliable) and those of the children (91% chances on lengths 5, 7, 9 cm., and 99% chances on lengths 7 and 9 cm., for true differences). Like the children, the subnormal girls did particularly poorly on the shortest standard, 5 cm.,

[33] Correlation between the TT and TV scores was negligible (+0.09 ± .09), so that the P.E.$_{diff.}$ was computed assuming zero correlation between the two performances.

by both methods (TT av. 20.4 $\pm$ 6.75 and TV av. 36.48 $\pm$ 15.09), so that in the comparisons on lengths 7 and 9 cm. (length 5 cm. excluded), they performed in a manner decidedly superior to the children by the TT method and decidedly inferior to the children by the TV method.

The subnormals were more variable by the TV method than by the TT method (TT 85% as variable as TV), in the same degree as were the other two groups.

The comparisons for the three modes of procedure, L-R or R-L, initial TT or TV presentation, and longer-shorter or shorter-longer adjustment, showed that in some respects the subnormal group reacted more like the college students, in other respects more like the children. The quantitative records for these three procedures, respectively, were as follows:

(1) Left-right, right-left procedure. Standards 5, 7, 9, and 11 cm. No. of scores, 100

Method	Av. score	P.E.	Diff.	$P.E._{diff.}$	Crit. R.
LR	15.37	4.39	1.02	.64	2.7
RL	17.08	5.41			

These differences do not reach complete statistical reliability. It is merely indicated that the subnormals like the college students were influenced by the more usual direction of movement of the preferred hand (see Table III).[34]

(2) TT or TV presented initially. Standards 5, 7 and 9 cm. No. of Scores=75.

Method	Av. score	P.E.	Diff.	$P.E._{diff.}$	Crit. R.
TT_1	17.368	5.75	.998	.88	1.1
TT_2	16.37	5.05			
TV_1	26.61	11.28	2.01	1.71	1.2
TV_2	28.62	9.25			

There are no reliable differences between the average scores. As we suggested previously, the degree of 'awareness' of the nature of the difficulties in the tasks would cause a difference in the behavior of sophisticated and naïve *Ss* (see Table IV).

(3) LS or SL adjustments[35]

Method	Standards (cm.)	Adjust.	No. of trials	Av. score	P.E.	Diff.	$P.E._{diff.}$	Crit. R.
TT	5, 7	LS	152	21.85	11.57	5.97	1.1	5.5
TT	5, 7	SL	148	15.88	6.98			
TT	9, 11	LS	148	13.69	5.77	4.93	.73	6.8
TT	9, 11	SL	152	18.62	6.9			
TV	5	LS	98	41.94	18.72	12.71	2.6	4.9
TV	5	SL	52	29.23	12.91			
TV	9	LS	76	24.0	9.98	3.16	1.6	1.9
TV	9	SL	74	20.84	9.18			

[34] Left-handed *Ss* were eliminated from this group, as in Experiment I.
[35] See footnote 20 for method of computing scores here.

The differences in the average scores are completely reliable except in the case of length 9 cm. by the TV method. The subnormals, hence, made larger errors where the longer adjustment was required (with one exception), as did the children (see Table V). They reacted like the college students on length 9 cm. by the visual method. This was the easier task, as the differences in length here for the LS and SL adjustments were less than on length 5 cm.

For the subnormals the tactual-kinesthetic reproductions were oftener shorter than the standards, while their visual reproductions measured were oftener longer than the standards, as was the case among the college students and children. These frequencies in percentage for the 4 standards are:

	Greater	Less	Equal
TT	40	56	02
TV	72	27	01

By the TT method, the ratio of frequencies 'smaller to greater' than the standards is only 7:5, instead of 2:1 as in the case of the other two groups. On the shortest standard, 5 cm., the subnormals overestimated the reproductions to the same extent as the children did (frequencies greater than the standard, 61%, less than the standard, 35%). By the TV method, on the other hand, the ratio of frequencies 'greater to smaller' than the standards becomes 8:3, showing a larger proportion of 'greater frequencies' than in the case of the college students, and still larger than in the case of the children (see Table VII, rows opposite 'mean').

The percentage of measurements 'greater' or 'less' than the standard lengths by the subnormals were influenced by the direction of the adjustments, LS or SL, mainly as they were in the case of the children and college students (Table VII shows percentages for the three groups). In one respect the subnormals reacted more like the adults than like the children: the percentages of measurements exceeded the standards by the SL as well as by the LS procedure in the TV method more than they fell short of the standards. For the subnormal group this tendency was very marked on length 5 cm.

In order to compare further the differences in performance of the three groups of *Ss* on the TT and TV methods we determined the percentage of the measurements exceeding or falling short of the standards by 2 cm. or more (Table X). By the TT method the three groups were most accurate on length 7 cm. On length 5 cm. the greatest number of exaggerations were made by the children, fewer by the subnormals, and very few by the adults. On the two longer extents, 9 and 11 cm., among all three groups a number of cases occurred in which the lengths were underesti-

mated 2 cm. or more. By the TV method the subnormal group made the greatest number of exaggerations on lengths 5, 7, and 9 cm. On length 11 cm. both adult groups exaggerated more frequently than did the children. On this length the children exaggerated slightly less than they underestimated.

TABLE X

PERCENTAGES OF ESTIMATIONS AT LEAST 2 CM. 'GREATER THAN' (>) AND 'LESS THAN' (<) STANDARDS

Method	Standard cm.	Children		Adults		Subnormals	
		>	<	>	<	>	<
TT	5	.23	0	.03	.007	.15	.01
	7	.09	.05	.04	.05	.08	.04
	9	.02	.21	.02	.11	.07	.16
	11	.04	.31	0	.23	.04	.28
TV	5	.35	.007	.21	.01	.38	.02
	7	.30	.05	.29	.05	.40	.03
	9	.27	.08	.29	.03	.45	.03
	11	.14	.17	.29	.03	.29	.09

The tendency among college students, and to an even greater degree among the more naïve adults, to interpret visual reproductions differently from tactual-kinesthetic reproductions of initial tactual-kinesthetic perceptions of extent, seems to be a form of perpetual illusion, and one more apparent among adults than among children. That this illusion is one of a space-time relationship seems to us to be the most plausible explanation for the following reasons.

Research on estimations of extent in different perceptual fields has shown the reciprocal influence of duration and distance.[36] The less the speed of presentation of a given extent, the longer is this extent judged to be.[37] Now in our experiment the time necessary for the initial movement of the index finger was longer than that for scanning the visual field when the visual reproductions were to be made. The total visual field was sufficiently small to be taken in 'at a glance,' that is, the eye-movements required in looking back and forth over a maximal distance of 13.5 cm. were few and could be carried out rapidly in comparison with the time consumed in moving the index finger across the same space interval at the rate of 10 cm. every 3 sec. No one as far as we know has considered this space-time relationship when a shift is made from perceptual functioning in one sensory field to that in another sensory field. There is no reason, however, why the illusion would not take place under these conditions. We feel, therefore, that the longer time necessary for perceiving to the full extent the initial mechanical stimulus and the shorter time required for perceiving the light-pattern stimulus account for the tendency to make the visual reproductions measure longer than the standards.

[36] For references of experiments on this problem in the field of vision, touch and kinesthesis, see J. Fröbes, *Lehrbuch der experimentellen Psychologie*, 1, 1923, 350, and H. Helson and S. M. King, The tau-effect: An example of psychological relativity, *J. Exper. Psychol.*, 14, 1931, 202-217.

[37] This illusion has been called the "tau-effect" by Helson, *op. cit.*, 203.

The subnormal girls were most under the influence of this illusion and the college students less so, possibly because of their awareness of this effect, and their consequent attempt to compensate for it. The children were least influenced by the illusion.

In addition, we noticed a mode of behavior characteristic of the children that possibly counteracted the space-time illusion. A great many of them when making visual adjustments went through the motion of moving the index finger of their right hand either in front of the exposure-screen or on their laps. In other words, the children attempted to duplicate the initial movement at the same time that they were observing the visual field and, consequently, obtained more direct kinesthetic cues in making their judgments than did the adults and subnormals. Among the college students only 3, and none in the subnormal group, made overt finger movements of this kind that we could detect. If they made incipient or tentative movements they naturally obtained cues from the same source as did the children, but their cues were more tentative and reduced; though none the less important for them.

It may be that another illusion was present that accounted for the discrepancies between the visual and tactual-kinesthetic perceptions. In a given sense-modality, filled space is usually apprehended as longer than unfilled space.[38] There is no reason to think that this effect could not carry over from one perceptive field to another, from the tactual to the visual. It is a matter, of course, for careful investigation. We merely suggest here that in our experiment the continuous tactual perception, as the index finger moved across the edge of the pasteboard, might have given the impression of a filled tactual space, whereas the visual perception of hollow rectangle, bounded on three sides by gray and on the far side by black cloth might have given the impression of empty space.

Children react more or less like adults in having the more usual perceptual illusions within one sensory field.[39] It may be that in the shift from touch to vision, the perceptual integration among children is not sufficient to produce an illusion effect.

The reduced tactual-kinesthetic reproductions have received various explanations.[40] It is possible that the space-time illusion operates here also. We do not know the speed of the movements made while the *S*s of our

[38] For an experiment on filled and unfilled space illusions in the field of touch, see F. B. Dresslar, Studies in the psychology of touch, this JOURNAL, 6, 1893, 332-342.

[39] A. Binet, La mesure des illusions visuelles, *Rev. philos.*, 40, 1895, 11-25; Dresslar, *op. cit.*, section on Size-weight illusion in children, 343-357; and R. Meili and E. Tobler, Les mouvements stroboscopiques chez les enfants, *Arch. de psychol.*, 23, 1931, 131-156.

[40] J. Loeb, Untersuchungen über die Orientirung im Fühlraum der Hand und in Blickraum, *Pflüger's Arch. f. d. ges. Physiol.*, 46, 1890, 1-46; C. O. Weber, The properties of space and time in kinaesthetic fields of force, this JOURNAL, 38, 1927, 597-606. For further references, see Weber, footnotes, 603-604.

experiment carried out their tactual-kinesthetic reproductions. If they tended to be slower and more cautious during the adjustment than during the initial perception, then the 'reduced' reproductions are accounted for by the influence of the longer time required to traverse a given extent. Whatever the rate of movement, it undoubtedly was not uniform.

Possibly the explanation lies in the further influence of load or difficulty. In a study on perceptions of extent of movement, Weber stipulated that when distances are reproduced under load, all *Ss* tend to give a distance shorter than the unloaded standard.[41] Kramer and Moskiewicz had previously suggested that when movements were troublesome or awkward (movement of less preferred hand) they were judged longer than easy or smooth movements across the same extent.[42] It may be that the greater difficulty of having to make a decision during the adjustment process produced, among our three groups of *Ss*, the illusion of greater extent or load. At any rate, this illusion was not as great as the illusion during the visual reproductions.

We have seen from the results of this experiment that subnormal girls who have reached physical maturity performed better than the children by the TT method. We anticipated this result, because of the fact that the subnormals had had more years of training and practice in coördinated hand and arm movements. They did less well than the superior adults, lacking no doubt the ability to make finer discriminations, particularly on length 5 cm.

By the TV method, the subnormals did less well than both the children and the college students. Dr. S. T. Orton has suggested in conversation with the writer that the reason for this poor showing in the visual reproductions as compared with the tactual-kinesthetic reproductions might be that mentally retarded individuals are particularly deficient in the development of their visual association areas. Their tactual-kinesthetic association areas, on the other hand, function more adequately and in a more stable manner, due to their longer phylogenetic history. Consequently, Dr. Orton would expect them to perform more easily by our TT than by our TV method, and that their showing by the latter method would be particularly poor, as compared to that of individuals whose visual association areas were more developed. This explanation does not seem to be entirely justified in view of our results. In the first place, the superior adults made a rather poor record by the TV method; in fact, their average score on the two middle standards, 7 and 9 cm., was not reliably better than that of the children. In the second place, if the visual association areas of the defectives were not developed to a degree necessary for carrying out the visual reproductions of tactual-kinesthetic perceptions of extent, then the errors made by the subnormal girls would be more likely random errors, the frequencies of measurements exceeding the standards approximating the frequencies of those falling short of the standards. We found rather that the errors by the TV method among all the three groups lay predominantly in one direction, that of positive deviation, above the standards. This trend was most pronounced among the

[41] Weber, *op. cit.*, 599.
[42] F. Kramer and G. Moskiewicz, *Zsch. f. Psychol.*, 25, 1901, 121.

subnormals and the college students. It would follow then that although the *Ss* lacked adequate cues for making correct estimates, they nevertheless had sufficient cues for integrating visual and tactual-kinesthetic mechanisms, to the extent, at least, that a perceptual illusion occurred between the initial perception and its visual reproduction. As we have suggested, this illusion may have been produced by a space-time relationship or by the effect of filled and empty space.[43]

General Discussion

It would be important, of course, before drawing final conclusions about the relationship of tactual-kinesthetic and visual mechanisms of children and adults, to repeat our experiment under conditions where the procedure is reversed: to present visual extents and require the *Ss* to reproduce these extents by visual means in one method and by tactual-kinesthetic means in the other method. Nevertheless, we shall discuss our results as far as they go, taking into account the limitations set up by the conditions of the immediate experiment.

As has been seen, the proficiency of the children in making reproductions of tactual-kinesthetic perceptions of extent in the tactual-kinesthetic field was not noticeably superior to their proficiency in making visual reproductions of the same extents. The skill of the two adult groups, college students and subnormals, on the other hand, was considerably greater when they made tactual-kinesthetic reproductions of tactual-kinesthetic perceptions, than when they shifted to vision as the dominant means of control, and made visual reproductions of the same tactual-kinesthetic perceptions.

These results appear to be contrary to those of Renshaw, where the task required was cutaneous localizations by both tactual-kinesthetic and visual means.[44] In his experiment, as may be recalled, the children performed better than the adults when they employed tactual-kinesthetic mechanisms, while the adults performed more efficiently with predominantly visual mechanisms.

The differences in results need some explanation. In the first place, the perception of short horizontal movements of the right arm, with tactual cues obtained from the tip of the index finger, and kinesthetic cues from various muscle groups of the shoulder, arm, and hand, is a more practiced

[43] We obviously do not question the fact that retarded individuals have undeveloped visual association areas, which will show up in tasks requiring finer or more complicated discriminations than those needed in carrying out the visual reproduction task in this experiment.

[44] S. Renshaw, The errors of cutaneous localization and the effect of practice on localizing movement in children and adults, *Ped. Sem.*, 38, 1930, 223-238; Renshaw and Wherry, Studies on cutaneous localization, III, The age of onset of ocular dominance, *ibid.*, 39, 1931, 493-496.

mode of functioning than that of perceiving and localizing stimulated areas on the skin-surface. In the course of their daily lives, individuals spend many more hours in handling, manipulating, and moving objects, writing and drawing, than in locating points of stimulation on the body surface. It is comprehensible, therefore, that the performance of the adults in our experiment should have been superior to that of the children when they made use of tactual-kinesthetic mechanisms.

Lack of practice alone, however, does not explain why the adults did less well than the children in cutaneous orientations by tactual-kinesthetic means. As we have previously stated, Renshaw formed the hypothesis, based on neurological evidence, that the tactual-kinesthetic mechanisms of adults have suffered a decrement because of the appearance of vision as the dominant means of sensory control.[45] But this explanation does not hold true for the results of our experiment, where the adults did decidedly less well when they employed visual mechanisms than when they made use of predominantly tactual-kinesthetic mechanisms.

As we indicated, the TV task in our experiment was a much less practiced one than the TT task for the adults, especially as the cues had to be obtained from the less usual order of sensory control, from touch to vision, rather than from vision to touch. But the adults in Renshaw's experiment had no more practice for carrying out the TV task, where they had to indicate on an outline sketch of the hand or arm the point at which they had been stimulated on the skin. What then is the explanation for the differences between our results and those of Renshaw when adults made visual estimates?

We think a partial explanation lies in the fact that the *Ss* in our experiment, particularly the two adult groups, were influenced in their performance by perceptual illusions, of space-time or filled versus empty space, whereas the *Ss* in Renshaw's experiment did not have the same illusions because the stimulation of the skin was instantaneous and a point rather than a linear contact. In other words, the presence of the illusions may have prevented the adults in our experiment from functioning relatively as well as the adults in the experiment on cutaneous localization.

We still feel, however, that Renshaw's hypothesis of vision as the domi-

[45] As further proof for his point of view, Renshaw pointed out that blind adults, for whom vision could not be a deterrent to the smooth running of tactual-kinesthetic mechanisms, made skin localizations by the TT method more accurately than did seeing adults; their performance being comparable to that of seeing children. Renshaw, Wherry, and Newlin, Cutaneous localization in congenitally blind versus seeing children and adults, *Ped. Sem.*, 38, 1930, 246-247. In our experiment, on the other hand, where the practice effect was greater, blind adults should make tactual-kinesthetic reproductions comparable to those of seeing adults.

nant sensory control among adults is correct and is borne out by the results of our experiment. We suggest that the reason for the more predominant illusions among our adult *Ss*, as compared with the children, is the presence of this visual dominance. We can say then that ocular dominance aided the performance of the adults in one type of visual task (cutaneous localization), but interfered with their performance in another type of visual task (reproduction of tactual-kinesthetic perceptions of extent).[46] Further, it is our judgment that in the one task, cutaneous localization, the unpracticed tactual-kinesthetic mechanisms functioned less well because of the dominance of the visual mechanisms, while in the other, perception of extent, the more practiced tactual-kinesthetic activity offset any influence from the dominant visual field.

Another factor that could account for the differences in results might be that the visual task in Renshaw's experiment was harder for the children than our task of making visual reproductions. The identification of a visual sketch of hand or arm with the actual hand or arm involves more complicated functioning (imagination) than the perceptions of extents where both the mechanical energy for the tactual perception and light patterns for the visual perception originate from the same stimulus source.[47] Under conditions where the visual task is less complicated and more comparable to the tactual-kinesthetic task, it may well be that children are relatively at a greater advantage than adults in making visual reproductions, not only because of the fact that adults are essentially no more practiced than are children in this mode of activity, nor that they are more under the influence of visual illusions, but because of a more inherent difference: adults have greater differentiation between the tactual and visual modes of functioning, more specificity in touch discrimination and in visual reproductions of tactual perceptions. Children, on the other hand, have more global or "syncretic" impressions of space, regardless of the sense-modality in which these perceptions are initiated.[48]

Other investigators, who have observed the development of perception in the young child, have pointed out the large rôle played by tactual factors in the perception of visual form, and that, in all perceptions of children, the various sensory

[46] Dr. Orton agreed that visual imagery may interfere with a kinesthetic task. He mentioned a case he had who could draw a geometric figure from memory provided he did not look at what he was doing.

[47] In a first experiment on two 12-yr.-old boys, however, Renshaw employed a procedure where the visual estimates were made on the skin. The *Ss* fixated the cross-sectional square printed on the back of the hand, and called out to *E* the coördinates of the spot localized as the one that had been stimulated (*op. cit.*, 234).

[48] E. Claparède, Exemple de perception syncrétique chez un enfant, *Arch. de psychol.*, 7, 1908, 195-198.

fields (particularly touch and vision) blend more readily into one indiscriminate total experience than they do among adults.[49]

The perceptual activity then, on the part of the children, was not capable of making the finer integrations required in this experiment between tactual and visual sense-modalities. The perceptions of the adults, on the other hand due particularly to their more dominant visual sensory control, were more synthetic. The two adult groups, normal and subnormal, made attempts, even though crude and subject to illusory effects, to integrate visual reproductions of extent with previously apprehended tactual-kinesthetic extents, to a greater degree than did the children.

On the basis of his results showing the superiority of children in making localizations by tactual-kinesthetic means, Renshaw suggested that children may, in our present mode of education, be provided with too many purely visual methods of instruction, and that they should be allowed to make more use of their contact and kinesthetic mechanisms.[50] It would seem on the basis of our results that as the sense-modalities of touch and vision are less differentiated in the child, more use could be made of their contact and kinesthetic mechanisms, though not at the expense of their visual mechanisms.

Furthermore, on the basis of our results, it is indicated that subnormals who had reached a fair degree of physical maturity reacted perceptually in many respects more like normal adults than like children. It remains for further investigation to determine how far, in other directions, the perceptual functioning of adult defectives parallels that of normal adults.

---

[49] Cf. H. Volkelt, Fortschritte der experimentellen Kinderpsychologie, *Ber. ü. d. IX Kong. f. exper. Psychol.*, München, 1926, 80-135, R. Meili, Les perceptions des enfants et la psychologie de la Gestalt, *Arch. de psychol.*, 23, 1931, 25-44, and S. Szuman, Geneza przedmiotu o dynamicznej integracji sfer zmyslowych we wczesnem dzieciństwie (Genesis of "objects": Concerning dynamic integration in the sensory fields in early childhood), *Kwartal. Psychol.*, 3, 1932, 363-394.

[50] Renshaw, Wherry, and Newlin, *op. cit.*, 247.

THE LOCALIZATION OF ACTUAL SOURCES OF SOUND

By S. S. Stevens and E. B. Newman, Harvard University

Quantitative information relative to the ability of an *O* to localize actual sources of sound in free space is prerequisite to the formulation of an adequate theory of localization. The acquisition of such information depends upon the possibility of generating pure tones for a wide range of frequencies and of presenting them to an *O* in such a way that no reflected waves reach the ears. The present study attempts to satisfy these two conditions.

Antecedents

The earliest systematic investigations of localization were conducted by what might be called the 'sound-cage school.' Their results were limited by the fact that, owing to the absence of electrical generating apparatus, they were forced for the most part to use clicks and noises as stimuli. Another limiting factor was their custom of experimenting in closed rooms whose walls were not sound-absorbent. In spite of these drawbacks, certain facts were established.[1] *O*s are able (1) to locate noises better than tones and (2) to distinguish right from left. However, they tend (3) to confuse the location of sounds lying in the median plane and (4) to locate sounds at the sides with the least accuracy. The greater inaccuracy of localization at the sides is exactly what one would expect on the basis of recent measurements[2] of the variation of the difference in loudness at the two ears of a speech-source rotated in a horizontal plane around the head. Throughout an angle of about 70° on either side the difference in loudness at the two ears remains virtually constant. In the face of this constancy it is obvious that within the 70°-angle there could be no intensitive cues for exact localization. An area of inexact localization at the sides is also indicated as the effect of the difference of phase at the two ears,[3] when the stimulus is a tone of low frequency.

From the sound-cage attention turned to the matter of the presentation of sounds which differ at the two ears in respect of intensity, phase, time of arrival, or a combination of these factors. In this way there arose the 'dichotic school,' whose problem it was to determine the relative merits of the intensity-theory, the phase-theory, and the time-theory.[4] Each of the differential factors—intensity, phase and time—influences localization, each has been nominated by one or more experimenters as the most

* Accepted for publication October 1, 1934.

[1] A. H. Pierce, *Studies in Auditory and Visual Space Perception*, 1901, 52.

[2] J. C. Steinberg and W. B. Snow, Physical factors in auditory perspective, *Bell System Tech. J.*, 13, 1934, 247-260.

[3] G. W. Stewart, Phase relations in the acoustic shadows of rigid sphere, *Phys. Rev.*, 4, 1914, 252-258; R. V. L. Hartley, Function of phase difference in binaural localization of pure tones, *ibid.*, 13, 1919, 373-385.

[4] O. C. Trimble, The theory of sound localization: A restatement, *Psychol. Rev.*, 35, 1928, 515-523.

important factor in localization, and each has also been reduced in theory to one of the others. The facts established by this school are (1) that localization is towards the side of greatest intensity, (2) that in the case of low tones lateral localization can be got by advancing the phase at one ear, and (3) that localization is towards the side of the sound which leads in time. Presumably phase-difference is but a special case of time-difference[5] so that in reality there are just two factors available as cues

FIG. 1. THE EXPERIMENTAL SET-UP
Note the absence of vertical reflecting surfaces at the level of the observer.

for the localization of pure tones. A final answer as to the rôle of these two factors could not be given, however, since it has not been known to what extent actual sounds can be localized.

A novel attack on the problem of localization, only recently initiated, is the investigation of auditory perspective: the stereophonic effect of multiple sources of sound.[6] Sounds picked up in one room by three microphones spaced several feet apart and broadcast through three correspondingly placed loud speakers in another room can be localized with considerable accuracy. The localization is what would be expected on the basis of the intensitive difference at the ears of the *O* as calculated

[5] E. M. von Hornbostel and M. Wertheimer, Über die Wahrnehmung der Schallrichtung, *Sitzber. d. preuss. Akad. d. Wissensch.*, 1920, 388-396; E. G. Boring, Auditory theory with special reference to intensity, volume, and localization, this JOURNAL, 37, 1926, 157-188.

[6] H. Fletcher et al., Auditory perspective: A symposium, *Bell System Tech. J.*, 13, 1934, 239-310. The essential facts of perspective with two-channel transmission were demonstrated in the Harvard Psychological Laboratory in 1929; cf. unpublished affidavit by M. Upton and W. D. Turner.

from measurements[7] of the effect of the sound-shadow cast by the head. Up to the present time, however, such studies have dealt only with complex sounds.

Apparatus

The present experiments were conducted entirely in the open air (Fig. 1). In order to avoid possible reflecting surfaces a tall swivel chair was erected on top of a ventilator which rises 9 ft. above the roof of the new Biological Laboratories at Harvard University. *O* was thus placed in a position where there were no vertical reflecting surfaces on any side of him, and the nearest horizontal surface on the side toward the source of the sound was approximately 12 ft. below him. At the present time this procedure appears to be the only practicable means of avoiding errors due to the reflection of sound.

The source of the sound was mounted on the end of a 12-ft. arm attached to the pedestal of the chair. When properly counterbalanced, it could be moved noiselessly in a complete circle in a horizontal plane on the level of *O*'s ears.

For the larger portion of the experiments a small 4-in. magnetic speaker, mounted in a 12-in. baffle, served to generate the tones. Only in the case of the lowest frequency (60 cycles) was it necessary to use a large Western Electric type-560 cone speaker in order to obtain sufficient power. This speaker was mounted on the same arm at a distance at 6 ft. from *O*.

A beat-frequency oscillator, which could be adjusted to the desired frequency, supplied ample power to the loud speaker. The voltage was adjusted by means of a 7000-ohm potentiometer, while a further shunting resistance was used to turn the tone on and off without producing clicks. The elimination of clicks in the loud speaker is, indeed, an important consideration. The tones were made reasonably pure by the use of suitable filters: tones of 2200 cycles and above passed through a high-pass section with a cut-off frequency of 2500 cycles; tones 3000 cycles and below passed through suitable low-pass sections which reduced the partials other than the fundamental by at least 30 db. The 60-cycle voltage was obtained directly from the lighting circuit. Two check experiments with unfiltered power at frequencies of 400 and 1000 cycles showed that slightly more accurate localizations could be expected with these less pure tones.

Comparative results for two noises were obtained by the use of a click and a hiss. The click was produced by applying 45 volts from a battery to the loud speaker for a brief instant. It was heard by *O* as a single sharp click which possessed the high frequency characteristic of the speaker. The hiss was produced by blowing air through a small brass tube, the end of which had been cut and pinched. The brass tube was attached to the end of the swinging arm and was blown through a long rubber tube by *E*. The sound produced contained a perceptible high-pitched whistle with a frequency of about 7000 cycles. The remaining energy was probably distributed over a wide band of frequencies.

The sensation level of the tones used in this experiment was determined approximately by comparison of the voltages used with threshold voltages determined under comparable conditions in a soundproof and sound-deadened room. Comparison of the sensation levels thus determined with standard audibility curves shows that tones of the middle range from 400 to 4000 cycles had a loudness level of 50 to 60 db. Tones

[7] L. J. Sivian and S. D. White, Minimal audible sound fields, *J. Acous. Soc. Amer.*, 4, 1933, 288-321.

in the extreme ranges, although they were of moderately high intensity, reached a loudness level of about 30 db.

The apparent intensity of the tones was reduced in the experimental situation by the presence of a constant background of sound consisting of distant traffic noises. The level of this noise was determined by a comparison of the thresholds obtained in the soundproof room with those obtained on the roof. The masking effect is measured by the rise in threshold for the various tones. The average rise in threshold for five frequencies between 400 and 7000 cycles, the region in which the effect was greatest, was 28.0 db.

Procedure

In the present experiment each of the two authors acted alternately as *O* and *E*. The tones were presented at certain definite positions in the horizontal plane of the *O*'s head. The *O*'s task was to name the position of the source of sound in this plane. We found in some preliminary experiments conducted in the spring of 1933 that reversals of right and left practically never occurred. Therefore, in the main series of observations which were made during the summer of 1934, the sounds were presented at 13 different positions on the right side of the *O*. These positions were spaced 15° apart from 0° directly in front to 180° directly behind. Ten observations were made at each position by each *O*.

The *O* tried, of course, to distinguish sounds in front from sounds behind, but his success, as will be shown later, depended upon what sound was being used. Since front-back reversals were frequent, we decided that the fairest measure of localizability could be obtained only if such reversals were not counted as errors. Therefore, the size of the error in the localization of a given sound was obtained by taking the difference between the reported position and either the actual position or the corresponding position in the other quadrant, depending upon which was the smaller. Thus, when the source was at 0° the localization was considered correct if it were either 0° or 180°; if the source was at 30°, both 30° and 150° were considered correct responses. This procedure is equivalent to the assumption that dichotic differences of phase or intensity provide a basis for lateral localization alone.

Localization As a Function of Frequency

The average of the errors made by both *O*s at each of the frequencies used is shown in Fig. 2. The errors, computed by the method described above, are relatively constant at low frequencies, but become larger as the frequency approaches 3000 cycles. This change accords with the general findings of the 'dichotic school.' However, above 4000 cycles localization improves again and is fully as accurate at 10,000 as at 1000 cycles. This result, as far as we know, has not been anticipated by previous experimenters. In fact, it has been suggested[8] on several occasions that high tones can not be localized at all—for no better reason, it would seem, than that no one had really tried to localize high tones.

The explanation of the shape of the curve in Fig. 2 must concern us.

[8] H. M. Halverson, The upper limit of auditory localization, this Journal, 38, 1927, 97-106, esp. p. 97.

The inexact localization of tones between 2000 and 4000 cycles is precisely what we should expect from a consideration of the effects of the two localizing factors, difference in phase and in intensity. Owing to the size and shape of the head, there are certain theoretical limits to the possible effectiveness of each of these factors. The limits are shown graphically in Fig. 3. It is well established that differential phase is most effective in determining the localization of low tones, and that above about 600 cycles its effectiveness decreases with increasing frequency. In Fig. 3 are shown the results (dotted line) obtained by Halverson[9] for the maximal lateral shift in localization obtainable with 180° phase-difference as a function of frequency and also the results to be expected theoretically. The theoretical curve (solid line) is a first approximation and was obtained by considering the difference in the distance a sound-wave would have to travel to reach the two ears. If the radius of the head is taken as 8.75 cm. the difference in distance is given by

$$d = 8.75\ (\sin\Theta + \Theta).$$

Since our method of treating front-back reversals reduces the problem to a consideration of the phase-effect within a single quadrant, we should not expect confusions to arise until the frequency is so high that two or more positions within a quadrant would give rise to the same phase-difference at the ears of the *O*. The frequency at which this first occurs is that whose wave-length is equal to the maximum value of d, or 1520 cycles.

Fig. 3 also shows the difference in intensity at the two ears for tones of different frequency originating at the side of the *O* (dot-dash line). This curve is due to Steinberg and Snow[10] and shows that the difference at the two ears tends to increase with frequency, as one would expect in view of the sharper sound-shadows obtained with high frequencies.

It is obvious from Fig. 3 that at low frequencies the phase-effect is ample to account for localization, and that at high frequencies the intensity-effect is sufficiently marked to afford good cues. Furthermore, in the region of 3000 cycles neither phase nor intensity is available as a differential cue. Hence the sharp maximum in the curve in Fig. 2 at about this frequency. In general, then, we can conclude that at low frequencies the localization of pure tones is made on the basis of phase-differences, at high frequencies it is made on the basis of intensitive differences, and that in a region near 3000 cycles localization is poor, because of the absence of both types of differences.

Additional reason for believing that differences in intensity alone are

[9] Halverson, *loc. cit.*
[10] Steinberg and Snow, *loc. cit.*

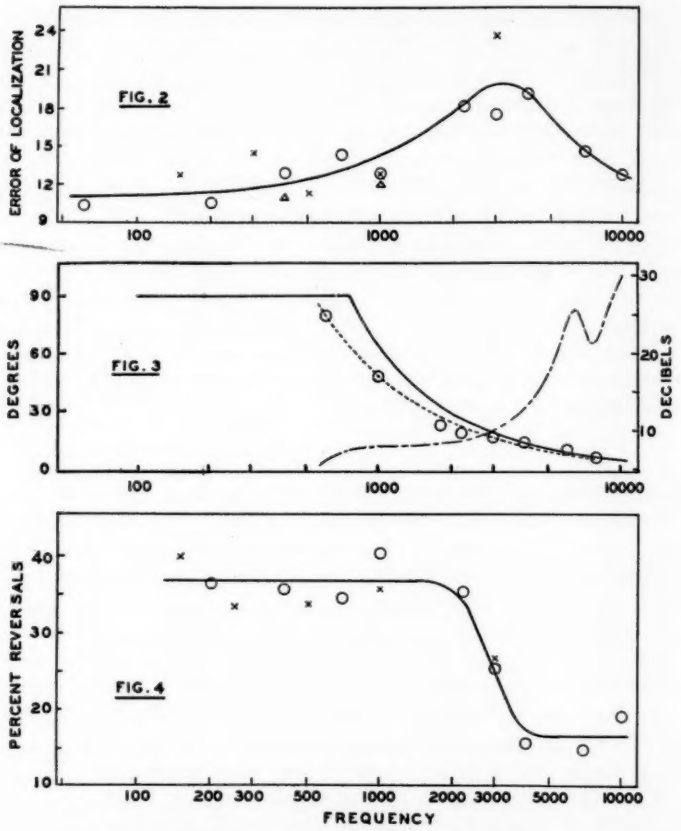


FIG. 2. DEPENDENCE OF LOCALIZATION ON FREQUENCY

The ordinate represents the average of the errors in degrees made by both *Os*. The crosses are for the shorter series of judgments made in 1933. The circles represent the results obtained in 1934. The triangles represent the results obtained with unfiltered tones. Note the critical region at about 3000 cycles.

FIG. 3. ABSENCE OF PHASE-EFFECT AT HIGH FREQUENCIES AND OF INTENSITY-EFFECT AT LOW FREQUENCIES

The solid curve represents theoretically the maximum angle by which a tone can be displaced by 180° change in phase. The circles on the dotted curve are the observed maxima of displacement (Halverson). The dot-dash curve represents the observed difference in intensity at the two ears of tones originating at the side of the observer (Sivian and White).

FIG. 4. PERCENTAGE REVERSALS OF THE FRONT-BACK QUADRANTS

The crosses are for the data obtained in 1933, the circles for 1934. The critical region is at about 3000 cycles; cf. Fig. 2.

operative at high frequencies is afforded by the fact that at about 2800 cycles the action currents in the auditory nerve cease to be synchronized with the stimulus.[11] This fact suggests that although below 2800 cycles phase-differences might well be effective since they could be transmitted as such to the brain, whereas above 2800 cycles any discrimination in terms of the phase-effect should be impossible. On the other hand, Halverson's results show at this frequency no sharp break which could be associated with a breakdown of synchronization. However, Halverson reports that marked changes of intensity occurred as the phase of the high-frequency tones was altered. It may be that his results were due to changes in intensity rather than phase.

Localization as a Function of Position

Since the work of Bloch[12] in 1893, it has been generally recognized without serious contradiction, that the localization of sounds in the horizontal plane is most accurate in the region directly in front or behind the *O*.

The average results for the fourteen main series of our experiment are shown in the following table.

Position in degrees from median plane	0°	15°	30°	45°	60°	75°	90°
Average error in degrees	4.6°	13.0°	15.6°	16.3°	16.2°	15.6°	16.0°

The values for 60° and 75° are weighted means. The weights were assigned in proportion as the errors due to reversals exceeded those expected on the basis of a normal distribution of errors. In order to determine these values the mean average deviation of the distribution was used: it is 15.8°. Using this value the number of reversals to be expected when the stimulus was at a given position could be determined from a table of the probability integral. The 'weighted mean' is thus a mean which includes those errors of localization in the other quadrant which would be expected as errors and which considers the remaining errors in the other quadrant as true reversals.

The results for individual frequencies show considerable variability because of the limited number of cases involved in each average. With the exception, however, of the three critical frequencies, 2200, 3000, and 4000 cycles, they are in substantial agreement with the table presented above.

---

[11] H. Davis, A. Forbes and A. J. Derbyshire, The recovery period of the auditory nerve and its significance for the theory of hearing, *Science*, 78, 1933, 552.

[12] E. Bloch, Das binaurale Hören, *Zsch. f. Augenheilkunde*, 24, 1893, 25-86.

Since localization at low frequencies is based on the phase-effect, we should expect the errors of localization to be inversely proportional to the rate of change of phase-difference as a function of azimuth, or equal to $k/(\cos\Theta + 1)$. The errors made by *N* at low frequencies agree fairly well with this expectation, but those made by *S* are too large at the 15° and 30° positions. We have reason to suspect a systematic error at these positions.

In general, the shape of the function relating errors and position is what might be expected upon theoretical grounds.[13] The fact that the function is essentially the same at the very high frequencies, however, requires explanation. The greater effectiveness of differential intensity near the median plane is not, as early investigators supposed, an example of Weber's Law. On the contrary, the effect can be predicted from the way in which the relative intensity at the two ears depends upon the azimuth of the sound. Both the theoretical values of Hartley and Fry[14] and the experimental determinations of thresholds by Sivian and White[15] indicate that a relatively large displacement of the source is necessary at the side of the *O* to produce a given change of intensity. It is to be expected, therefore, that localization would be best near the median plane and poorest at the extreme lateral positions.

Somewhat more striking in our experiments was the ability of the *O* to distinguish between front and back. The relative frequency of reversals has been taken as a measure of front-back discrimination. In Fig. 4 are presented the average results of both *O*s plotted against frequency. The individual curves agreed closely in form with the average curve. One function (for *N*) was displaced upward and slightly to the left of the other (for *S*).

It is apparent at once that the total range of frequencies is divided into two distinct regions separated by a narrow critical range at about 3000 cycles. For tones below 2000 cycles, where localization is based on phase-differences, discrimination between the front and back quadrants is only a little better than chance. Above 4000 cycles the number of reversals is but one-third of those expected by chance. Taken in connection with the other data presented in this paper, these results offer striking confirmation of the double mechanism (differential phase and intensity) involved in normal localization. The ability of the *O* to distinguish front from back in the case of tones of high frequency is very largely a function of the

[13] Stewart and Hartley, *loc. cit.*

[14] R. V. L. Hartley and T. C. Fry, The binaural localization of pure tones, *Phys. Rev.*, 18, 1921, 431-442.

[15] Sivian and White, *loc. cit.*

difference in intensity between sounds in front and behind.

A number of checks were made in order to demonstrate the validity of this hypothesis. A continuous tone of high frequency, when swung about the *O* decreased very markedly in loudness from front to back. A number of tests were made in which intensity was varied from trial to trial in chance order; and the percentage of reversals, for six series at 3000 to 7000 cycles, was found to increase from 18.6% with the usual procedure to 47% with this test procedure. It appears that the *O* had formed a subjective standard of intensity in the main series after a very few trials. After this the tones in back "seemed weak" while those in front were loud and close. Sound-shadows from the pinna are probably sufficient to account for this effect.

Localization as a Function of Complexity

The experimental series in which the click and the hiss were used confirm the fact that complex sounds are more easily localized than pure tones. The localization of the hiss in particular was very definite, almost as definite as though one were looking at the source! The average of the errors of localization was only 8.0° for the click and 5.6° for the hiss. Thus it may be remarked that the fact that the click was much better localized than any of the pure tones shows the importance of eliminating all of the clicks and extraneous noises which are present if the onset of a pure tone is too abrupt. In the case of the hiss neither *O* was aware of any intensitive or qualitative difference at the different positions during the test series; but careful observation afterwards revealed noticeable differences of both kinds. When in front the hiss was more *shh* and less *sss* and was louder than when behind. The qualitative differences were due, of course, to the differential effect of the shadow of the head upon the frequencies composing the hiss.

In view of findings of the present experiment relative to the ease with which high tones can be localized, it is not surprising that noises with high-frequency components are very easily localized. Noises in general have both high and low frequencies present. The low frequencies provide sufficient phase-differences and the high frequencies sufficient intensitive differences for localization. The two types of cue render each other mutual support, and the result is an accuracy of localization greater than that obtainable with pure tones.

Conclusions

(1) The ability to localize tones varies markedly with frequency. It is approximately constant below 1000 cycles, drops rapidly to a minimum

between 2000 and 4000 cycles, and rises again to its former level at higher frequencies. See Fig. 2.

(2) The error of localization is smallest for tones located near the median plane and increases as the tone is moved toward the side of the *O*. The relation between the azimuth of the tone and the error of localization is approximately the same for both high and low frequencies.

(3) The confusion of positions lying in the quadrant in front of the *O* with those in the quadrant behind him is very frequent. Below 3000 cycles the frequency of such reversals was about that which should be given by chance. Above 3000 cycles it was only about one third of the chance value.

(4) Noises (a click and a hiss) were localized more readily than any of the tones. Differences of quality and intensity were discernible between different positions of the noises.

(5) All of the above facts are consistent with the hypothesis that the localization of low tones is made on the basis of phase-differences at the two ears, and that the localization of high tones is made on the basis of intensitive differences. There is a band of intermediate frequencies near 3000 cycles in which neither phase nor intensity is very effective and in which localization is poorest.

THE ADAPTATION OF PAIN AROUSED BY COLD

By Barbara Edes and Karl M. Dallenbach, Cornell University

Pain was once regarded as a "general quality of sensation . . . produced in all cases where the stimulation of a sensory nerve passes a certain limit of intensity."[1] This view of pain persisted until near the turn of the last century when from the welter of fact and opinion[2] two new and opposing

* Accepted for publication May 25, 1935.

[1] Oswald Külpe, *Outlines of Psychology,* 1893, 90. Titchener held to this view of pain even as late as 1896; "excessive stimulation of any sense-organ, or direct injury to any sensory nerve, occasions the common sensation of pain" (*An Outline of Psychology,* 1896, 65). For the relation of this view of pain to the doctrine of the *Gemeingefühl,* cf., Johannes Müller, *Handbuch der Physiologie des Menschen,* 2, 1837, 494; E. H. Weber, Der Tastsinn und das Gemeingefühl, in Rudolph Wagner's *Handwörterbuch der Physiologie,* 3, (2), 1846, 466 f., 553-588. Weber's section was reprinted separately as *Die Lehre vom Tastsinne und Gemeingefühle,* 1851, 21 f., 108-143.

[2] E. Kröner, Gemeingefühl und sinnliches Gefühl, *Vjsch. f. wiss. Philos.,* 11, 1887, 153-176.

B. Naunyn, Über die Auslösung von Schmerzempfindung durch Summation sich zeitlich folgender sensibler Erregungen, *Arch. f. exper. Path. u. Pharmakol.,* 25. 1889, 272-305.

H. R. Marshall, The psychological classification of pleasure and pain, *Mind,* 14, 1889, 511-536.

A. Goldscheider, Ueber verlangsamte Leitung der Schmerzempfindung, *Dtsch. med. Woch.,* 16, 1890, 688-689.

Goldscheider, Ueber die Summation von Hautreizen, Arch. f. [Anat. u.] Physiol., 1891, 164-169.

H. R. Marshall, The physical basis of pleasure and pain, *Mind,* 16, 1891, 327-354, 470-497.

Herbert Nichols, The origin of pleasure and pain, *Philos. Rev.,* 1, 1892, 403-432, 518-534.

H. R. Marshall, Pleasure-pain and sensation, *idem,* 625-648.

Alexander Bain, Pleasure and pain, *Mind,* 17, 1892, 166-187.

Henry Head, On disturbances of sensation with especial reference to the pain of visceral disease, *Brain,* 16, 1893, 1-333; 17, 1894, 339-480.

H. R. Marshall, *Pain, Pleasure, and Aesthetics,* 1894, 3; Pleasure-pain, *Mind,* 19, 1894, 533-535.

Lightner Witmer, Pleasure and pain from the psychologist's standpoint, *Amer. Med. & Surg. Bull.,* 7, 1894, 351-353.

H. R. Marshall, Are there special nerves for pain? *J. Nerv. & Ment. Dis.,* 21, 1894, 71-84.

Lightner Witmer, The psychological analysis and physical basis of pleasure and pain, *idem,* 209-228.

C. A. Strong, The psychology of pain, *Psychol. Rev.,* 2, 1895, 329-347.

Herbert Nichols, Pain nerves, *idem,* 487-490.

H. R. Marshall, Physical pain, *idem,* 594-598.

C. A. Strong, Physical pain and pain nerves, *ibid.,* 3, 1896, 64-68.

Herbert Nichols, Pain nerves, *idem,* 309-313.

Emil Schiff, Der Schmerz, *Die Nation,* 13, 1896, 295-298.

Charles Richet, Étude biologique sur la douleur, *Rev. scient.,* 6, 1896, 225-232.

M. L. Frederico, Y a-t-il des nerfs spéciaux pour la douleur? *idem,* 713-715.

Charles Richet, Résponse, *idem,* 715-717.

theories—Von Frey's and Goldscheider's—emerged, engendered by the discovery of the specific and punctiform character of cold, warmth, and pressure.[3]

Both Goldscheider and Von Frey agree that pain is a specific sensation and that it is derived from a special class of nerves, but from here on their theories diverge. Goldscheider holds that pain is mediated by the tactile nerves and that it results from a summation of their excitations: weak stimulation gives touch; intense stimulation, pain.[4] Von Frey, on the contrary, holds that pain is an unique, separate modality of experience and that it is subserved by separate receptoral structures—the free nerve endings.[5] The difference between these two theories is, as Kiesow wrote in his critical review of them, "so deep that a reconciliation is hardly conceivable . . . if one view is correct the other must necessarily be false."[6]

The authors of these contradictory theories both point to the results of experiments on pain-adaptation for demonstration and support. Goldscheider claims that pain not only becomes weaker under continuous stimulation but that it also undergoes a qualitative change. Before coming to complete adaptation and disappearing it passes, as he believes, into a sensation that he characterizes as sub-painful *(unterschmerzliche Empfindung)*, tactual *(Berührungsempfindung)*, or pressure *(Druck)*. Von Frey claims in opposition, that pain simply becomes weaker and weaker with continued stimulation and finally disappears without undergoing any qualitative change whatsoever.

The problem of pain-adaptation has, in consequence of the disagreement between these two authors, become of particular significance. The special problem of adaptation is transcended by the more general implications of the results.

Adaptation experiments in which pain is aroused by mechanical stimuli,[7] either

---

[3] For a bibliography of these studies see K. M. Dallenbach, this JOURNAL, 41, 1929, 344.

[4] A. Goldscheider, *Über den Schmerz in physiologischer und klinischer Hinsicht*, 1894, 13 f.; *Gesammelte Abhandlungen*, 1, 1898, 44, 397; Über Schmerz und Schmerzbehandlung, *Zsch. f. physik. u. diät. Therap.*, 19, 1915, 289-309; *Das Schmerzproblem*, 1920, 20, 71; Beiträge zur Physiologie der Gemeingefühle; I. Physiologie des Schmerzes, *Zsch. f. Sinnesphysiol.*, 57, 1926, 1 ff.

[5] Max von Frey, Beiträge zur Physiologie der Schmerzsinnes. *Ber. ü. d. Verhandl. d. königl. sächs, Ges. d. Wiss. z. Leipzig*, math.-phys. Kl., 46, 1894, 185-196, 288-296; Beiträge zur Sinnesphysiologie der Haut, *ibid.*, 47, 1895, 166-184; *Untersuchungen über die Sinnesfunktionen der menschlichen Haut*, 1896, 242. Der Schmerzsinn, in M. von Frey and H. Rein's Physiologie der Haut, 1929, 142 (Vol. 1, pt. 2, of the *Handbuch der Haut- und Geschlechtskrankheiten*, 1929).

[6] Friedrich Kiesow, The problem of the conditions of arousal of the pure sensation of cutaneous pain, *J. Gen. Psychol.*, 1, 1928, 199.

[7] In addition to Von Frey's and Goldscheider's studies, the following experiments

sharp or blunt, have yielded results that in general corroborate Goldscheider; pain is replaced by pressure experiences before the stimulus becomes completely ineffective. Since these stimuli, as Burns and Dallenbach pointed out,[8] are always applied at superliminal pressure values, they arouse pressure as well as pain. The pressures appearing in the course of adaptation may, therefore, be nothing more nor less than the residual effects of the stimuli employed. The only way of deciding whether that is actually the case is "to arouse pain with stimuli, for example radiant heat, that do not give rise simultaneously to pressure experiences."[9] Von Frey attempted to do this with a sun-glass. He focused the rays upon his skin and obtained, as he thought, a pure pressureless pain that neither *unterschmerzlich beginnt* nor *unterschmerzlich aufhört.*[10] In refutation of that experiment and its results, Goldscheider pointed out that under intense stimulation the course of experience is telescoped and obscured; that intense stimuli, such as burning heat, are therefore unfavorable for the observation of the *unterschmerzliche Empfindungen;* and that the latter are to be observed only under two favorable conditions: (1) when the pain is not intense, and (2) when it rises gradually to its full intensity.[11]

When these conditions, the gradual onset and growth of pain, were met in Stone and Dallenbach's experiment with radiant heat,[12] the experience aroused not only became weaker with time but also underwent a qualitative change; *warmth* emerged before the experience completely disappeared.

Since a pressure-pain stimulus yields residual pressure on its way to complete adaptation, and a warmth-heat-pain stimulus yields residual warmth, Stone and Dallenbach concluded that the accompaniments of pain in its course of adaptation are functions of the type of stimulus employed in its arousal. They further hazarded the prediction—which set the problem of the present study—that if a cold-pain stimulus were employed, a residual cold-effect would be noted.

The purpose of the present study is, then, twofold. We wish (1) to discover whether the pain aroused by cold would come to complete adaptation under continuous stimulation; and, if so, (2) to determine the residual effects.

Method and Procedure

It was necessary for the purpose of this experiment to devise some means of arousing cold without simultaneously arousing pressure. Moreover, the cold had to be intense enough to produce pain and yet not intense enough to obscure any pressures that might otherwise be experienced. After many unsuccessful attempts, these requirements were met by using dry-ice (solidified carbon dioxide) and the stimulator shown in Fig. 1.

---

have been made: H. H. Straus and F. R. Uhlmann, Adaptation of superficial pain, this JOURNAL, 30, 1929, 422-424; E. F. Wells and L. B. Hoisington, Pain adaptation: A contribution to the Von Frey-Goldscheider controversy, *J. Gen. Psychol.*, 5, 1931, 352-357; M. Burns and K. M. Dallenbach, The adaptation of cutaneous pain, this JOURNAL, 45, 1933, 111-117.

[8] Burns and Dallenbach, *op. cit.*, 115.

[9] *Ibid.*, 116.

[10] Von Frey, Versuch über schmerzerregende Reize, *Zsch. f. Biol.*, 76, 1922, 9.

[11] Goldscheider, *op. cit.*, *Zsch. f. Sinnesphysiol.*, 57, 1926, 5.

[12] L. J. Stone and K. M. Dallenbach, Adaptation to the pain of radiant heat, this JOURNAL, 46, 1934, 235.

Apparatus. The stimulator, made of brass, consisted of three parts: (1) the body, a cylindrical reservoir for dry-ice, 8 cm. long and 3 cm. in diam.; (2) the cap, with tube and cross rod; and (3) a funnel-shaped nose-piece or nozzle with an opening 2 mm. in diam.[13]

A screen between the ice-chamber and nose-piece prevented small pieces of dry-ice from falling into and plugging the opening.

The stimulator was insulated with several layers of sheet asbestos so that the cold aroused derived only from the nozzle. The tube and cross rod of the cap served a double purpose: as a handle to remove the cap when the reservoir was

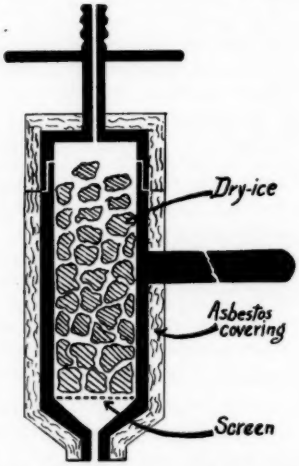


FIG. 1. DIAGRAM OF THE STIMULATOR

to be filled with dry-ice, and as an inlet for the compressed-air current that was used in the preliminary experiments. We at first used the stimulator, after the manner of Sullivan and Verda's apparatus,[14] with compressed air, the pressure of which was controlled by a manometer. We abandoned the use of compressed air after we found that when the nozzle of the apparatus was brought close to the skin, any pressure upon the air column was sufficient to arouse pressure experiences. Indeed, even without compression, enough air passed through the open tube of the stimulator to arouse a sensible pressure. Our problem of arousing cold without simultaneously arousing pressure was not solved until we packed the inlet tube with cotton wool, and permitted the stream of carbon dioxide to flow out of the bottom of the apparatus to the skin under the pressure that was built up by the evaporation of the dry-ice.

[13] This is the same diameter as the radiant heat stimulus used by Stone and Dallenbach, *op. cit.*, 232.

[14] A. H. Sullivan and D. J. Verda, The experimental fusion of warmth and cold in heat, *J. Exper. Psychol.*, 13, 1930, 208 ff.

Observers. The *O*s were Mr. W. S. Neff (*N*), assistant in psychology, and Miss S. G. Longwell (*L*), Miss M. L. Lemmon (*M*), and Miss L. K. Allen (*A*), graduate students in psychology. *N* had observed in an earlier experiment on adaptation to pain.[35] The other *O*s were untrained in cutaneous observation. All the *O*s, however, knew in general the purpose of the experiment.

Instructions. The following instructions were given the *O*s at the beginning of every observational period.

At a 'Ready, Now' signal your arm will be stimulated. Describe the experiences aroused and follow their temporal course. Remain as motionless as possible. Close your eyes at the signal and do not open them until told to do so.

Procedure. The shaved dorsal surfaces of both forearms were used as experimental tissues. *O* was seated in a chair beside a table with his arm in an arm-rest. At the 'ready' signal the stimulator was brought above the arm. At 'now' it was turned down by means of the rack-and-pinion gear of a universal support to a point 1 mm. from the skin and held there till 20 sec. after the last report of pain, when it was removed. Then, at intervals varying from 30 to 180 sec., it was reapplied to the same spot, unless, as frequently happened, after-sensations of pain had in the meantime been reported. The spots were restimulated in order to determine whether the disappearance of pain just reported was due to adaptation or to the destruction of the functional nerve endings.[36]

E recorded *O*'s reports and noted the corresponding times from a stop watch which was started as soon as the stimulator was brought to the skin.

A preliminary series of experiments of variable length served to familiarize the *O*s with the nature of their task. This was followed by the main series of 25 experiments.

Results

Normal course of experience. The following excerpts from the reports of the *O*s illustrate the normal and typical course of experience resulting from stimulation of the skin by radiant cold.

N, (Exp. 14): (0″) cold; (43″) slight pain; (55″) pain increasing slightly; (1′) very faint pain; (2′10″) very diffuse and not intense cold; (3′) pain more definite; (3′53″) pain quite definite and better localized, fading off into cold; (5′27″) certain amount of fluctuation of intensity; (7′45″) pain decreasing in intensity; (7′55″) occasional twinge of sharper pain; (8′10″) nothing but very faint cold; (8′30″ stimulus removed).

A, (Exp. 9): (0″) cool; (1′) cooler and more concentrated; (1′21″) aches slightly; (1′22″) burns very slightly and aches slightly; (2′56″) burn more concentrated, not intense; (3′46″) burn more piercing; (4′40″) burns and aches but not intensely; (4′44″) ache disappearing and burning not as intense; (5′5″) burning very faint; (5′25″) cool; (5′34″) faint burning; (6′1″) cool; (6′21″) stimulus removed).

L, (Exp. 21): (0″) cold; (13″) faint sting; (24″) sting with cold more intense on ulnar side; (48″) sting increasing slightly; (2′2″) sting decreasing, but cold is as intense; (2′57″) sting has changed to dull ache; (3′38″) ache almost disappeared; (4′7″) slight sting; (4′12″) ache; (5′) just cold; (5′4″) slight sting; (5′6″) just cold; (5′26″ stimulus removed).

M, (Exp. 11): (0″) cool (2″) nothing; (5″) cold; (20″) prickly and cold; (2′33″) decidedly sharp pricking; (4′5″) only faint pricking; (4′27″) dull ache

[35] Stone and Dallenbach, *op. cit.*, 233.

[36] Cf. Stone and Dallenbach, *op. cit.*, 240.

and cold; (4′58″) just cold; (5′3″) faintly cool; (5′13″) nothing; (5′18″ stimulus removed).

Almost immediately upon stimulation every *O* reported "cool" or "cold," which increased in intensity and passed over into pain. The pain, in turn, grew in intensity, remaining at a high degree for a time, then decreased and cold again became prominent. Eventually, the pain was entirely displaced by cold. In some of the experiments the cold faded out completely.

The experiences aroused by radiant cold may be divided, as those aroused by radiant heat, into a fore-phase, a phase beginning with the application of the stimulus and ending with arousal of pain; a mid-phase, extending from the appearance of pain through its growth to its decline; and an end-phase, running from the decline of pain to removal of the stimulus.

Aside from this typical course, there were a few variant types. Sometimes along with the usual experiences of pain there were sharper intermittent twinges. In this case, pain often adapted for a few seconds, only to return before the stimulus was removed. The following report offers an example of this.

N, (Exp. 18): (0″) cold; (30″) cold with slight bit of pain; (47″) pain more intense; cold less evident; (1′2″) occasional twinges of sharp pain overlaying continuous pain; (1′55″) twinges of sharp pain; continuing pain of average intensity; (2′6″) more cold coming in; (2′30″) cold increasing; pain present; (3′25″) pain less intense; (3′36″) cold well defined; (4′3″) certain amount of throbbing under pain area; still a good deal of cold; (4′40″) occasional twinge of sharper pain comes and goes almost immediately; (5′25″) has character of painful cold; (6′0″) just cold; (6′12″) pain; (6′15″) pain gone; (6′30″) pain returned; (6′40″) pain gone; (7′0″) still very faint pain; (7′10″) just cold; (7′30″ stimulus removed).

Arousal of pain. After the stimulus was applied to the arm, some time elapsed before pain was aroused. As shown in Table I, the average time ranged from 66.2 $\pm$ 36.6 sec. (for *M*) to 31.5 $\pm$ 15.2 (for *A*). The individual times for the different *O*s ranged from 6 to 225 sec.

TABLE I

SHOWING THE AVERAGE TIME (IN SEC.) REQUIRED FOR THE AROUSAL OF PAIN, THE M.V., AND THE RANGE OF DISPERSION

O	Av.	m.v.	Range
A	31.5	15.2	6–80
L	59.7	31.9	10–225
M	66.2	36.6	12–225
N	53.2	24.2	7–132

In addition to its slow arousal, pain was, as the following characterizations of the first pains experienced clearly show, fairly mild and weak.

A: aches slightly (18); aches (4); slight searing sensation (2); dull ache (1).

L: stings slightly (9); stings (5); aches (3); dull ache (2); slight ache (2); pain (1); itchy sting (1); slight itch (1); slight pain (1).

M: faint pricking (9); pricking (4); pain (2); sore (2); piercing pain (1);

painful cold (1); slight ache (1); prickly and cold (1); dull ache (1); slight prick (1); slight pain (1); stinging cold (1).

N: slight pain (17); pain in center of an area of cold (5); beginnings of pain (2); very, very faint pain (1).

The numbers in parentheses indicate the frequency of use.

Since these reports show that our stimulus, like the stimulus employed by Stone and Dallenbach, was not of a kind to produce suddenly an intense experience of pain, we too have met Goldscheider's conditions for the observation of the *unterschmerzliche Empfindungen.*

Pressury experiences. Since pain was neither abruptly nor intensely aroused, pressury experiences should have, according to Goldscheider's theory, been numerous. Only one *O,* however, reported pressure during the fore-phase of the experiment, and this *O* (*M*) did not relate it in any way with the pains subsequently experienced.

During the mid-phase, pressure was reported 4 times; but all of these pressures were present simultaneously with pain, and they were reported before pain had begun its decline. *N* reported three incidences of pain with a "drawing character," and one with pressure.

In the end-phase four reports of pressure were again given. *N* gave 2 reports referring to a "pain with a drawing, pulling character," and *M* reported a "drawn feeling" and a "pressury sensation."

In all, pressure was reported in 9 of the 100 experiments. Certainly these results afford no convincing evidence that pain is due to a summation of pressure excitations. Five of these 9 cases, moreover, occurred along with pain, and concomitance of pressure and pain tells against, and not for, summation. Of the remaining 4 cases of pressure only one occurred before pain was aroused.

Adaptation to pain. As stated before, 20 sec. was our criterion for pain-adaptation. This particular time-interval was taken so as to bring our results into line with those of previous studies. The interval was sufficiently long to insure that adaptation had really occurred and that *O*'s report was not due to a shift of attention.

In spite of the fact that our criterion was rather rigorous, complete adaptation was reported by all our *O*s in all the experiments. According to Table II, the average time for the adaptation of pain from its arousal first to its disappearance ranged from 185.6 $\pm$ 68.7 sec. (for *A*) to 317.8 $\pm$ 113.0 sec. (for *N*). The range of dispersion was found to be from 36 to 787 sec. The deviations may in part represent individual differences, but are more likely, as we believe, to have been caused by chance variations in the number, sensitivity, and tuning of the cold spots stimulated.

TABLE II

SHOWING THE AVERAGE TIME (IN SEC.) REQUIRED FOR THE ADAPTATION OF PAIN AFTER ITS FIRST AROUSAL, THE M.V., AND THE RANGE OF DISPERSION

O	Av.	m.v.	Range
A	185.6	68.7	69–396
L	295.5	85.8	41–643
M	188.4	92.1	36–438
N	317.8	113.0	70–787

In order to discover whether the disappearance of pain was due to adaptation or to the destruction of the endorgan, the stimulator, it will be recalled, was returned to the spot stimulated at various intervals (30, 60, 90, 120, and 180 sec.) following its removal. Table III shows that the time of

TABLE III

SHOWING THE AVERAGE TIME (IN SEC.) REQUIRED FOR THE REAROUSAL OF PAIN, THE M.V., AND RANGE OF DISPERSION, TOGETHER WITH THE DIFFERENCE BETWEEN THE TIME OF AROUSAL AND REAROUSAL

O	Av.	m.v.	Range	Difference between arousal and rearousal
A	54.5	30.3	8–135	22.9
L	64.3	50.6	2–225	4.6
M	68.1	45.4	2–235	1.9
N	44.7	26.4	12–145	−8.5

rearousal ranged from 2 to 235 sec. For all *O*s except *N* this time was longer than the original time of arousal (22.9 sec. longer for *A*, 4.6 for *L*, 1.9 for *M*). *N*'s time of rearousal was shorter by 8.5 sec. than the original time of arousal. In general, therefore, we may state that the 'pain-time' at restimulation is longer than at the original stimulation.[17] The results, however, were obtained from too few cases and the differences are too irregular and small to be conclusive. Their significance for us lies in the fact that pain reappeared in all but six cases, thus insuring that its disappearance was a matter of adaptation and was not caused by a destruction of the endorgan.

The end-phase consisted of a gradual decline in intensity of pain, which finally was replaced by cold. Sudden large drops in the intensity of the pain did not occur, nor was there anything to indicate masking or telescoping.

[17] Stone and Dallenbach explained the shorter period of rearousal in their experiment by a "softening effect" caused by the original warm stimulus. The longer rearousal period in this experiment might be due to a "deadening" effect of cold.

Summary

The results of this study may be summarized as follows: (1) pure, pressureless pain may be aroused by a cold stimulus not in direct contact with the skin; (2) pain is completely adaptable; (3) pain disappears without being followed by a residual pressure; (4) the disappearance of pain under continued stimulation by cold is a true case of adaptation and not a result of injury to the endorgan; (5) cold always precedes and usually follows pain when the latter quality is aroused by a cold stimulus; (6) the time of rearousal of pain is usually greater than the time required for its first arousal.

Our results confirm the view that (1) Goldscheider's theory of pain as a summation of pressures is untenable; (2) pain is adaptable; (3) pain is a separate modality; and (4) the accompaniments of pain—pressure, warmth, or cold—are functions of the kind of stimulus used in the arousal of pain.

THE INFLUENCE OF THE REPORTING-RESPONSE UPON THE REPORT

By S. M. NEWHALL and E. H. RODNICK, Yale University

To what extent are the responses by means of which a report is made, instrumental in determining the report itself? Psychologists have long been concerned with the dependence of the subject's report on present and past stimuli and past responses, but relatively little attention has been given to the dependence of the report on the present reporting-responses. Anyone would agree that responses made previously in the subject's life-time may play a large part in shaping the present report, but often the assumption is implied that the present response has little or no effect. Perhaps one feels that the report must have been devised before it could be expressed and that therefore the reporting-response is merely a conveyance or medium of expression. When an observer reports his introspection, for instance, he must have decided what he was going to say before he could say it. Necessarily, it would seem, the formation of a judgment precedes the response which expresses it;[1] and therefore it is the report which determines the response, not the reverse.

Still, *a priori,* it is not unreasonable to suppose that the response can affect the report. As a graphic record of a motor response is sometimes demonstrably invalid because the recording device has interfered with the response, so a subject's expression would be invalidated if the reporting-response itself was a faulty conveyor of the report. Alteration in the report might occur during its expression or delivery, or else the original data to be reported might themselves be changed; in either case the end result would be changed. The report which is modified might itself be of a relatively simple or complex nature. The effect might depend in large measure on various aspects of the total situation in which the subject reacts. In brief, the participation of the recorded response in the effect might be relatively direct or indirect, relatively large or small. The point could not be demonstrated directly because no report can be made known in the absence of a reporting-response; however, it would be demonstrated indirectly if different types of reporting-response yielded significantly different reports under adequately controlled conditions. The term 'might' has been used repeatedly since the whole matter indicates experimentation.

As a matter of fact, observation and experiment provide instances in which the response does seem to affect the report. A few illustrations will serve to indicate the complexity and pervasiveness of the general problem and to suggest some interesting experimental approaches.

Speed vs. accuracy. In sensory situations, a direct relation has been demonstrated between the accuracy of a judgment and the speed with which it is reported. The common explanation is that promptness of report is a consequence of feelings of certainty or confidence and the confidence correlates with accuracy. There seem, however, to be some situations in which this common 'cause-effect' sequence is reversed, that is to say, a voluntary speeding up of the reporting-response may improve the

* Accepted for publication June 5, 1934.

[1] H. L. Hollingworth, The definition of judgment, *Psychol. Rev.,* 32, 1925, 342.

quality of the report itself. Henmon found some evidence of this in discrimination of length of lines.[2] In various psychophysical experiments a relatively prompt report has been found to be the more reliable.[3] Experience with academic 'true-false' tests points in the same direction; too much consideration of a question not infrequently leads to an unfavorable reversal of judgment. Judging from the literature the consensus is that in relatively simple repetitive situations involving uncertainty, prompt response is best. The usual explanation seems to be that the first tentative judgment in the doubtful situation is likely to be best and that prompt expression of this judgment prevents vacillation. But when the report concerns some complex or difficult intellectual problem, speeding up the response may be conducive to error. In various forms of artistic or creative activity a calculated or leisurely response seems to work best. A quick response might interrupt the proper development of the report to be expressed.

Facility vs. precision. Difficulty of the reporting-response may operate to slow down the response. If the *S* in a psychophysical experiment, for instance, had to make some lengthy adjustment to signal his report he might have time to change his mind or he might become fatigued and less precise. Physicists and others who rely on the method of adjustment have long believed that an easy, comfortable adjusting response favors higher precision.

If the response is distracting, because of difficulty or some other reason, and tends to divert attention from the data to be reported on, it may thereby modify the data and so, indirectly, the report. An *S* starting a threshold experiment, for instance, may miss some of the first faint stimuli because of preoccupation with an unfamiliar mechanism for signalling his report. So he says he did not see the light nor hear the sound whereas he might have seen or heard it if his reporting response had been easier to execute.

Difficulty in response is likely to give rise to tension. Freeman observed that tension in an *S* changed the sensory threshold for shock.[4] If a reporting response involved tension, it might, therefore, indirectly modify the report by temporarily raising the limen. Interesting to recall in this connection is the fact that one of Spencer's *S*s, (*G*), was under marked tension and showed a definitely higher limen than the others.[5]

Skill and practice. In many instances, apparently, distortion of the report can be explained in terms of lack of skill, practice, steadiness, or aptitude. This may be the case whenever the reporting response consists in the reproduction of a pattern. Other things equal, the more complicated the pattern the less perfect may be the copy. Even though the *S*s know exactly what they want to do, and are ready to give a perfect report the expression may be imperfect because of the limitations in the reacting mechanism. The *O*s in a blind-spot experiment, for example, were requested to reproduce by sketching what they saw when the light stimulus was on the edge

[2] V. A. C. Henmon, The relation of the time of a judgment to its accuracy, *Psychol. Rev.*, 18, 1911, 188.

[3] Cf. *e.g.* S. M. Newhall and H. M. Halverson, Eye movements correlated with innervation of the orbicularis oculi, *J. Gen. Psychol.*, 11, 1934, 290.

[4] G. L. Freeman, The facilitating and inhibiting effects of muscular tension upon performance, this JOURNAL, 45, 1933, 17.

[5] L. T. Spencer, The validity of Heymans' law, this JOURNAL, 36, 1925, 430.

of the blind-spot. Presumably those skilled in drawing made the more accurate reproductions. Conceivably, all of them might have had the same or nearly the same perceptual experience in the experiment, but the investigator could not tell that from the sketches. The individual differences in the reporting responses might result in false conclusions with respect to individual differences in perception. Or, if all were unpracticed in drawing, the report of no one of them might reveal some crucial datum actually present in all cases.

Science and the law have long recognized the fact that imperfections in verbal reports are often due to lack of skill in that mode of expression. In a word, the report may be expected to vary with what Thorndike has called the *availability* of the proper response.[6]

Memory and association. Facts and theory of memory, association, and conditioning support the view that in the given situation some forms of reporting-responses favor recall more than others. The familiar rule to learn or practice in the way which one wishes ultimately to perform is aimed at the establishment of effective cues. The kinaesthetic sense supplies a mechanism by which habitual motor responses can take on recall value as well as other factors in the habitual situation. Incidentally, the mnemonic device to recall by writing as compared with trying to say it, is sometimes effective. When one form of reporting response is successful in a situation where others fail, that particular response or set of responses has made the report possible. In a sense, the response has effected if not affected the report.

Reserve and embarrassment. In an association experiment in which the *S*s were presented with lists of words with the instruction to report orally what each word brought to mind, some or all of them made reservations because of embarrassment or social pressure. If instead of reporting orally to the woman investigator, the *S*s had been allowed to write out their associations anonymously, the obtained data would presumably have been more complete. On the other hand, if the *S*'s task were to report his unfavorable opinion of some absent person he might prefer to state it orally rather than leave a permanent record by writing. These examples suggest the rather obvious fact that an inappropriate reporting response may result at times in a consciously curtailed or falsified report.

Mode of response. The fact that certain of the sense-modalities are intrinsically more sensitive and discriminating than others has long been appreciated. The fact of individual differences in sensitivity and differences in different animals is also common knowledge. Now analogous questions may be raised in respect of reporting responses. In an examination situation, for instance, some students are known to prefer the oral and others the written mode of report. Do such preferences possibly reflect some intrinsic or native difference in relative adequacy of these two major forms of reporting response? A consideration of extreme cases will show at once, that some native response mechanisms are intrinsically better suited to the expression of report in some situations than are others; usually the hand is better adapted than the foot to the expression of a complicated symbolic report. But where there are only two or three judgment-categories and the *S* only has to signal the chosen one by a simple movement, we do not know that the obtained data would vary with the response mechanism employed.

[6] E. L. Thorndike, *Human Learning*, 1931, 89-92.

Several investigators have touched upon the latter problem. Thomson wrote that "the subject communicated her judgments *one, two,* or *doubtful* to the clerk by signs. . . . The writer has not, however, detected any differences between those series of experiments in which this practice was followed and those in which the judgments were communicated by word of mouth."[7] In this study the signs were used to prevent *E* from knowing what *S* reported; the problem of the influence of different forms of report was not necessarily recognized.

Warner Brown explicitly recognized and experimented on the problem. "Even the very slightest differences in the task which he [the *S*] has to perform seem to govern to some extent his decisions." In one series his *S* had to report 'which' of the two stimuli was the heavier by indicating with the finger. In another series the procedure was to state orally whether the second weight was heavier or lighter than the first. The oral procedure gave the more accurate results. Unfortunately these two series differed in more than one respect. In the oral series judgment always referred to the second stimulus, whereas in the manual it might refer to either stimulus; the first was the more modern, the second the traditional Fechnerian procedure. The author inclined to the view that it was the 'pure Fechnerian element' rather than any motor element in the manual series which reduced the accuracy.[8]

Sears and Newhall experimented on the absolute visual threshold, using the method of constant stimuli. Data were secured simultaneously by two methods of response, viz., voluntary finger signal and conditioned finger retraction. The observed frequencies and the psychophysical thresholds calculated from both sets of data proved to be of the same order of magnitude indicating that the two modes of reporting response employed were about equally sensitive under the conditions of the experiment.[9]

Requisite force. The difficulty of a reporting-response sometimes is largely a matter of the physical force required to execute it. Would it make any difference, for example, if the *S* signalled his judgment by closing a key requiring 100-gm. pressure as compared with one requiring only 10? Would an *S* report just as faithfully if he had to turn a crank 5 revolutions to do so as compared with one revolution? At what point does the principle of least effort begin to make itself felt by distorting results?

Not infrequently an *S* is required to signal judgments in different categories by devices which have not been equated with respect to operating force. Whether small or moderate differences in requisite force would modify results is not known. The same is true for absolute force.

The ways in which some characteristic or concomitant of the response might possibly affect the report may now be summarized. The report might be reversed by too deliberate a response as in a true-false test; the precision might be lowered by a slow or inconvenient response as in a psychophysical experiment; the report might be distorted or even omitted because of inaptitude, distraction, or lack of

[7] G. H. Thomson, On changes in the spatial threshold during a sitting, *Brit. J. Psychol.*, 6, 1913-14, 434.

[8] Warner Brown, The judgment of difference, *Univ. of Calif. Pub.*, 1, 1910, 41-51, esp. 49 f.

[9] S. M. Newhall and R. R. Sears, Conditioning finger retraction to visual stimuli near the absolute threshold, *Comp. Psychol. Monog.*, 9, 1933 (no. 43), 21-23.

practice; failure to recall might result from employment of an unassociated response; the report might be consciously falsified because of embarrassment due to the prescription of an inappropriate response; *S* might fail to signal because of too much force required to execute the movement or, because of something of the nature of punishment. In brief, there are a number of potential sources of experimental error which may be considered from the point of view of the reporting-response.

A general solution of the problem is not feasible because of the multiplicity of conditions involved. The comparability of reporting-responses should be considered separately in every experimental inquiry. The technical task is to determine whether the potential errors are disturbingly large in the given experimental situation.

Our experiments were restricted to the psychophysical field for several reasons. First, experiments on the sensory threshold include relatively simple situations and relatively precise techniques; there exists, therefore, the probability of accurate measurements and a good chance of disclosing small differences. Secondly, the situation seems peculiarly sensitive in that awareness of very weak stimuli can be upset by relatively slight changes in conditions. Thirdly, the problem seems comparatively pressing here (a) because of the frequent importance of even small sources of error and (b) because of the diversity of reporting responses actually used in psychophysical investigations. As a rule, if an *S* can not speak because of head-rest or biting-board he uses a hand-key, or if his hands are occupied a foot-key may be employed. The assumption seems to be that the results will remain unaltered. Also, the possible effect of small or moderate differences in requisite force of the signalling reaction seems to be largely unrecognized.

We chose to work on the differential limen for brightness because it seemed most typical of discrimination work in general. But we were interested, too, in securing a certain check on the validity of the prior study by Sears and one of the writers.[10] In that study conditioned response and voluntary response had been compared in respect of sensitivity as threshold indicators. As suggested then, if it were found that a manual response were more sensitive than an oral response, some of the results of the study would need to be reinterpreted.

The present experiments, then, are confined to two questions: (1) what are the relative effects of oral, manual, and pedal reporting responses on the differential brightness threshold; and (2) what is the influence on the obtained threshold value of variation in the force required to execute the reporting response.

Influence of Mode of Response

The *S*s, 6 in number, faced a 3° photometric field in a dark room and they made their judgments of relative brightness by the traditional Fechnerian procedure. An artificial pupil was used throughout. Every *S* served in 3 series of observations: one in which he expressed his judgments by word of mouth; one in which he used hand-keys; and one in which he used foot-keys. Every one of them served in 6 sessions, and the series were temporally counterbalanced to permit comparison of the limens secured from the several series.

The limens for 'greater' and for 'less,' and their errors, were computed by the abbreviated constant process which eliminates the Müller-Urban weights. These values are presented in Tables I and II. They are in terms of Troland's photon and

[10] *Op. cit.*, 19.

the values of the means are seen to be almost exactly the same for oral, manual, and pedal response. Assuming the formula for the probable error of the limen to be approximately correct, the small differences present in case of the individual *Ss* are well within the range of random sampling. Therefore, we conclude that under the experimental conditions it did not make any substantial difference whether the *Ss* reported orally, manually, or pedally.

TABLE I

LIMENS FOR "GREATER" JUDGMENTS WITH ORAL, MANUAL, AND PEDAL REPORT

S	Limens			P.E.		
	Oral	Manual	Pedal	Oral	Manual	Pedal
1	50.89	51.22	50.57	.31	.34	.30
2	51.20	50.39	50.68	.28	.17	.26
3	50.55	50.96	51.24	.37	.39	.38
4	50.75	51.42	51.43	.26	.31	.33
5	50.96	50.58	50.75	.26	.31	.27
6	50.60	50.49	50.61	.22	.22	.21
Av.	50.83	50.84	50.88	.28	.29	.29

TABLE II

LIMENS FOR "LESS" JUDGMENTS WITH ORAL, MANUAL, AND PEDAL REPORT

S	Limens			P.E.		
	Oral	Manual	Pedal	Oral	Manual	Pedal
1	48.57	48.81	49.15	.30	.24	.29
2	48.92	48.75	48.22	.31	.30	.30
3	48.60	48.82	49.07	.34	.34	.30
4	48.77	49.22	49.10	.28	.26	.33
5	49.02	48.84	49.04	.25	.28	.21
6	49.74	49.29	49.52	.29	.24	.21
Av.	48.94	48.96	49.02	.30	.28	.27

The same conclusion follows from Fig. 1 which is constructed from the data of Tables I and II. The ordinate is discrete and refers to the type of reporting response used. The abscissa gives the unit in which the limens are calculated. Each circle represents the limen of some particular *S* as derived from 200 judgments. The solid circles or black dots are mean limens. Circles at the left represent limens for the 'less' judgments; circles at the right for the 'greater' judgments. The horizontal distance between any two circles represents the interval of uncertainty, half of which is often taken as the measure of the differential limen. The particular dot in the upper left represents, for instance, a limen for 'less' judgments of 48.57 photons; the corresponding threshold, at the right, for 'greater' judgments is 50.89 photons; and the interval of uncertainty is the difference, 2.32 photons.

INFLUENCE OF FORCE OF RESPONSE

Instrumentation was the same as in the other experiment except that manual keys only were employed and they were adjustible in respect of requisite closing pressure. The closing force for one or the other of these hand keys was always 10 gms.

whereas the other was either 100, 500, or 1000 gm. Therefore, the ratio of the one force to the other was either 10, 50, or 100. For purposes of this discussion, closing the 10-gm. key may be called the easy response and closing any one of the others the difficult response. As a matter of fact, however, the 'difficult' responses were

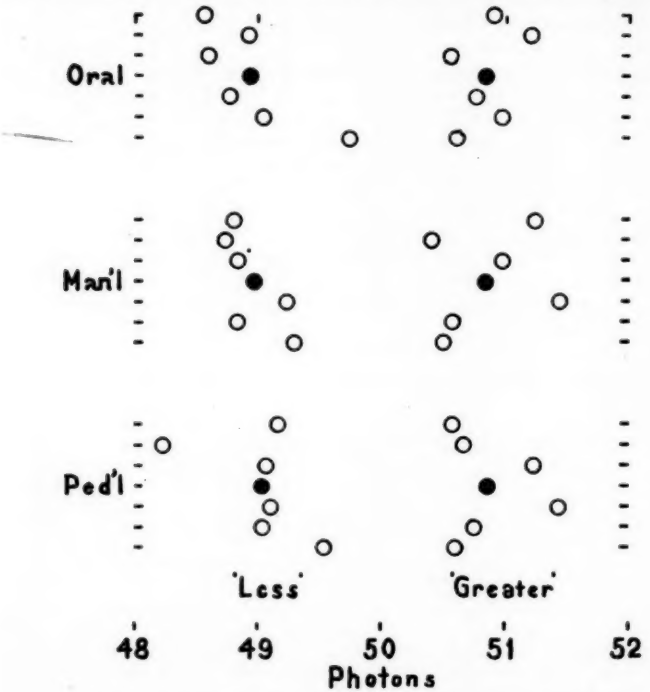


FIG. 1. LIMENS BASED ON ORAL, MANUAL, AND PEDAL REPORT

not really difficult. This was intentional as we wished to avoid here the moral issues in shirking an arduous task, and also to acquire data which might have some bearing on actual psychophysical practice.

The point, of course, was to see whether the easy response would result in a lower limen relative to the difficult response. The hypothesis seemed plausible that in cases of doubt or vacillation of judgment *S* might favor the easier response even though that response were not much easier. The principle of least effort might distort results even without *S*'s awareness. The *S*s were told nothing of the purpose of the experiment or of differences in closing pressure. They did not remark and did not appear to notice the smallest difference (10-100) in pressures. The other two (10-500, and 10-1000) were quite evident but they caused the *S*s no particular

concern except in the case of one *S* who reported the apparatus was defective and suggested that it should be repaired.

The results are exhibited in Tables III and IV and Fig. 2, which are arranged in a manner as analogous as possible to Tables I and II and Fig. 1. In Tables III

TABLE III

LIMENS FOR "GREATER" JUDGMENTS WHEN MORE FORCE IS REQUIRED TO SIGNAL "GREATER" COMPARED WITH THOSE WHEN MORE FORCE IS REQUIRED TO SIGNAL "LESS"

S	Force ratio	Limens			P.E.		
		G_m	L_m	$G_m - L_m$	G_m	L_m	$G_m - L_m$
1	10:1	50.77	50.43	.34	.24	.32	−.08
2	10:1	51.05	50.66	.39	.27	.43	−.16
3	50:1	50.76	50.23	.53	.19	.27	−.08
4	50:1	51.17	50.47	.70	.31	.29	.02
5	100:1	51.70	50.67	1.03	.31	.39	−.08
6	100:1	51.18	50.44	.74	.35	.41	−.06
Av.		51.11	50.48	.62	.28	.35	−.07

TABLE IV

LIMENS FOR "LESS" JUDGMENTS WHEN MORE FORCE IS REQUIRED TO SIGNAL "GREATER" COMPARED WITH THOSE WHEN MORE FORCE IS REQUIRED TO SIGNAL "LESS"

S	Force ratio	Limens			P.E.		
		G_m	L_m	$G_m - L_m$	G_m	L_m	$G_m - L_m$
1	10:1	49.68	49.49	.19	.30	.23	.07
2	10:1	49.67	48.96	.71	.32	.31	.01
3	50:1	49.63	49.03	.60	.25	.29	−.04
4	50:1	48.93	48.91	.02	.44	.41	.03
5	100:1	49.78	48.92	.86	.27	.29	−.02
6	100:1	49.79	48.77	1.02	.27	.29	−.02
Av.		49.58	49.01	.57	.31	.30	.01

and IV, the liminal values below the heading G_m are those secured when more force was required to signal 'greater' than 'less'; and those below the heading L_m, when more force was required to signal 'less' than 'greater.' The next column, that headed G_m-L_m, shows the differences in the values of the limens as due to the differences in force of response. While most of these differences taken separately are not significant statistically, the fact that there are no reversals by itself makes the results as a whole quite significant. In each of the 12 instances the evident effect of the physical resistance to response has been to increase the interval of uncertainty and reduce the sensitivity of report. Furthermore, as the force ratio rises the limens based on the difficult responses rise still more. This tendency for the effect to vary with the force-ratio confirms the view that it really is the force factor which is here active.[11]

[11] Morgan demonstrated a tendency to adjust effort to difficulty of task which Dallenbach believes is dependent upon or at least correlated with a change in the *O*'s attitude. Such a compensatory tendency if active in our *S*s would operate

In Fig. 2 the limens for the 'less' judgments are again at the left, for the 'greater' judgments at the right. The inscribed crosses represent the limens based on the relatively difficult reporting responses. These are connected by lines with the corresponding threshold based on the easy responses. From inspection it is evident that

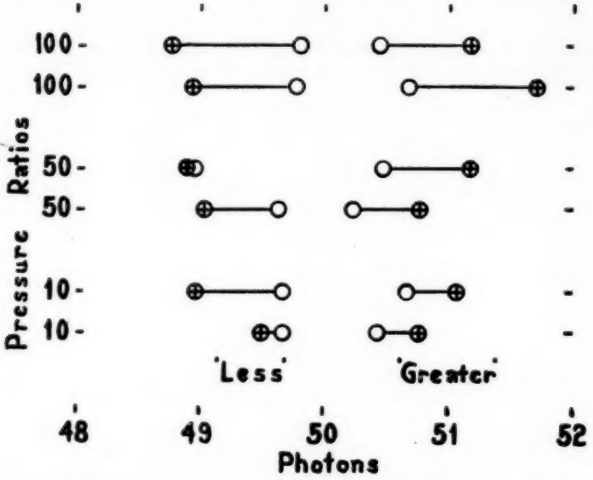


FIG. 2. LIMENS BASED ON MANUAL REPORT WITH VARIATION IN REQUISITE FORCE

the 'greater' thresholds based on difficult response are farther above the standard at about 50 photons than are the 'greater' thresholds based on easy response. Similarly, the 'less' thresholds based on difficult response are farther below the standard at 50 photons than are the 'less' thresholds based on easy response.

SUMMARY

(1) The problem of the effect of the response on the report appears to extend throughout the range of experimental psychology. Among the various response factors or concomitants which might singly or in combination influence the report in a given situation, may be included speed, facility, practice, association, embarrassment, modality, and force.

(2) The extent to which such factors operate as sources of significant experimental error by invalidating the *S*'s report is an important question which should, perhaps, be raised in connection with each experimental situation.

(3) In the field of psychophysics, experimentation on the differential brightness threshold with oral, manual, and pedal reporting responses, showed no essential differences in results. Thus some experimental justification was provided for the

to reduce distortion of report. Cf. J. J. B. Morgan, The overcoming of distraction and other resistances, *Arch. Psychol.*, 5, 1916 (no. 35), 80-84; and E. E. Cassel and K. M. Dallenbach, The effect of auditory distraction upon the sensory reaction, this JOURNAL, 29, 1918, 138-143.

frequent assumption in that field that reporting responses of different modality yield comparable results.

(4) Experimentation on the differential threshold with variation in the requisite force to signal the judgment, indicated that the so-called principle of least effort may operate unintentially or unconsciously. Even moderate differences in the force required to express the judgments apparently played a significant part in the determination of the obtained data. Relative force or resistance appears to be an important source of experimental error.

(5) The concept of optimum reporting response may prove useful in experimental psychology. An optimum reporting response might be defined as one which permits minimum distortion of the report by operation of disturbing factors in the given situation. The optimum response may be expected to vary to some extent with such factors as the experimental problem, procedure, or subject.

COLOR SENSATIONS PRODUCED BY INTERMITTENT SPECTRAL STIMULI

By GLENN A. FRY, Washington University

In a previous paper were described the color sensations produced by the intermittent exposure of a small bright white area.[1] The effects of varying the brightness of the intermittently exposed area, the duration and frequency of exposures, and the brightness of the surrounding area were investigated. The present paper deals with the effects of varying the wave-length composition of the stimulus, and special attention is paid to the effects obtained with low frequencies.[2]

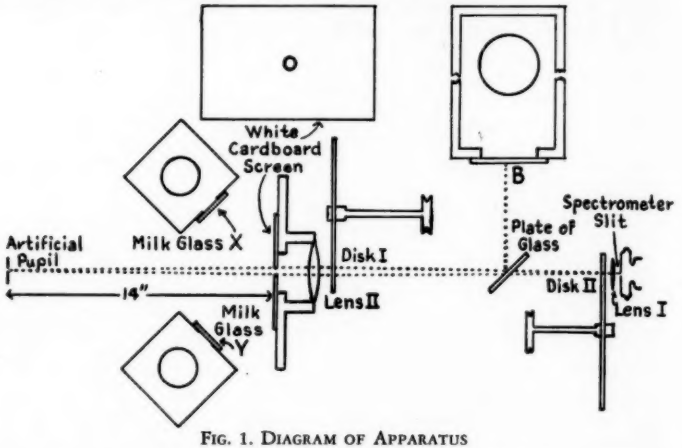


FIG. 1. DIAGRAM OF APPARATUS

The effects obtained with low frequencies (2 to 4 per sec.) differ from those obtained with higher frequencies in this respect, that with slow rates bright flashes alternate with dark intervals, corresponding to the on-and-off of stimulation, whereas with faster rates the on-and-off phases cannot be discriminated, and, although a fine flicker still obtains, one can assign to the experience a single brilliance and hue value. As the rate is increased from slow to fast the successive bright flashes may gradually fuse and thus blot out the dark intervals, provided the area is bright enough and the exposures long enough; but if, on the other hand, the stimuli are short and weak, the bright flashes may fail to appear and leave a dark color which increases in brilliance as the rate increases.

* Accepted for publication June 1, 1934.

[1] G. A. Fry, Modulation of the optic nerve-current as a basis for color-vision, this JOURNAL, 45, 1933, 488-492.

[2] The present paper, published by the author as National Research Fellow in the Biological Sciences, is from the Department of Ophthalmology, Oscar Johnson Institute, Washington University, St. Louis, Mo.

At an intermediate rate a peculiar phenomenon of rivalry often occurs. Periods of brilliant flickering alternate with longer periods in which a dark value obtains. A similar kind of rivalry occurs when sectors of red and green are alternated on the color wheel at a speed which is too fast for separate vision of the sectors and too slow for fusion into an intermediate color. First one and then the other predominates.

Effects with intermittent spectral stimuli were observed with the apparatus shown in Fig. 1. Light emerging from the spectrometer slit passed through Lens I and Lens II and through the round hole (½ in. diam.) in the white cardboard screen, and then formed an inverted image of the spectrometer slit (0.50 mm. wide), which fell vertically across a horizontal slit (0.64 mm. wide and 1.95 mm. long) used for an artificial pupil. The white cardboard screen was illuminated by light sources X and Y. The purity of the spectral light could be reduced by adding to it white light from source B, in order to show the transition to the effects obtained with white light. This white light from source B was reflected by the plate of glass placed at a 45°-angle in the beam of spectral light. The light coming through the hole in the screen was interrupted by Disk I. The on-off ratio was kept constant at 1/44, or, in other words, the size of the open sector was 8°. The duration of exposure varied from 5.4σ at 4 revolutions per sec. to 1.8σ at 12 revolutions per sec. The intensity of the light coming intermittently through the hole was kept constant at 2332 photons. It was impossible to obtain a pure blue (460 mμ) of this intensity, so one of 10% purity was used. The intensity of the surrounding white field was kept constant at 33.4 photons. The center of the hole was fixated. The results obtained with different frequencies of intermittent spectral stimuli are given in Table I.

When the rate is 4 per sec., a dark blue violet or purple interval is seen between the bright flashes in the case of white and all the spectral stimuli except red, and in the case of red also if the purity is reduced by admixture of white light.

The secondary purple images following the bright flashes are, perhaps, related to Bidwell's ghost. Bidwell projected an image of the spectrum (52.5 cm. long and 1.5 cm. wide) upon a white screen and caused it to be revolved parallel to itself in a circle 1 m. in diam. The observing distance was not stated, but a point near one end of the horizontal diameter was fixated. The rovolving spectrum was followed by a purple secondary image, which was shortened at the red end. This secondary image was separated from the primary image by a dark space which represented a time interval of about 1/5 sec.[3]

Nichols observed a purple secondary image of a stationary spectrum which was intermittently exposed with an on-off ratio of 1/35 and a very slow rate. Successive brilliant flashes, in which the characteristic colors of the spectrum were recognizable, were separated by dark intervals in which a uniformly purple secondary image of the spectrum was seen. The writer repeated this experiment using an on-off ratio of 1/71 and a rate of 2 per sec. and confirmed Nichol's finding except that the purple secondary image was shortened at the red end of the spectrum.[4]

Certain facts, however, differentiate Bidwell's ghost from the purple secondary

[3] S. Bidwell, On the recurrent images following visual impressions *Proc. Roy. Soc.*, 56, 1894, 132-145.

[4] E. L. Nichols, On the duration of color impressions upon the retina, *Amer. J. Sci.*, 28, 1884, 234-252.

TABLE I

EFFECT OF VARYING FREQUENCY AND WAVE-LENGTH OF STIMULATION

Wave-length	Color for different frequencies of stimulation			
	4 per sec.	6 per sec.	8 per sec.	10 per sec.
White	dark violet interval white flash	blue	blue green	greenish white‡
660 mμ	black* interval red flash	green† interval red flash	red	red
620 mμ	dark violet interval orange flash	dark blue	orange	orange
580 mμ	dark violet interval greenish yellow flash	alternating periods of greenish yellow and dark blue violet	yellowish green	greenish yellow§
540 mμ	dark purple interval green flash	green	green	green
500 mμ	dark purple interval green flash	green	green	green
460 mμ	dark purple interval blue flash	blue violet	blue	blue

* When the purity is reduced to 60 or 80%, the dark intervals acquire a purplish tinge.
† When the purity is gradually reduced the dark green changes to blue.
‡ The greenish white changes to white at 15 per sec.
§ The greenish yellow changes to yellow at 12 per sec.

images reported in the first part of this paper, inasmuch as the latter are dependent upon both intermittence of stimulation and the brightness of the surrounding field. When the field is brightly illuminated purple is not seen if the dark intervals are lengthened by allowing the flashes to come, say, once every 2 sec. In other words, single momentary flashes do not suffice; a repetition of the stimulus at a minimum frequency of 2 to 4 per sec. is necessary. On the other hand, if the rate is kept at 4 per sec. and the surrounding field is made dark, the purple is not seen, for under these conditions the hole is continuously filled with a bright flicker, the dark intervals being obliterated. However, if bright flashes occur only every 2 sec. while the field is kept dark, purple images are obtained. These are typical Bidwell's ghosts. Still it is not clear to what extent these purple sensations are related to those obtained with a higher rate of intermittence and with bright surroundings.

Since the stimulus used in the present investigation was a disk of ½ in. diam. at a distance of 14 in., thus subtending a visual angle of 2° 3′, it fell within the rod-free area of the retina; hence it may be assumed that cones alone were dealt with. When the periphery of the retina is used, as in Bidwell's experiment, rod-vision may also come into play; and when the stimulus intensity is weak, the secondary image may be mediated by rods alone. McDougall pointed out three facts which indicate

this fact. (1) The secondary image is not obtained with the fovea. (2) Red light fails to evoke it. (3) It is colorless regardless of the color of the primary image. When the intensity is raised cone-vision becomes involved, because red light produces a secondary image, and because it is colored and obtainable with the fovea. A red moving object, if bright enough, gives a red secondary image; a green moving object, a purple one; and a blue moving object, a blue one.[5]

Bidwell, in accounting for his ghost on the three-component theory of color-vision, postulated that the violet ghost was mediated by the violet component alone. He explained that red excites no violet ghost because it does not stimulate the violet visual component. It is perhaps wise not to try to explain the purple sensations reported in this paper because the rôles played by illumination of the surrounding field and intermittence of stimulation make the physiological problem complicated.

TABLE II

EFFECTS OBTAINED WHEN THE HOLE IS ILLUMINATED WITH STEADY WHITE LIGHT AND INTERMITTENT ADMIXTURES OF SPECTRAL LIGHT

Wave-length	Frequency of admixtures of spectral light	
	4 per sec.	6 per sec.
660 mμ	red flash green interval	red flash green interval
620 mμ	orange flash bluish gray interval	orange flash blue interval
580 mμ	greenish yellow flash puplish gray interval	yellowish white
540 mμ	green flash purplish gray interval	greenish white
500 mμ	green flash gray interval	green flash reddish purple interval
460 mμ	blue flash grayish yellow interval	purplish white

Also to be considered is the effect of steadily illuminating the hole in the screen with white light and then admixing to this the intermittent flashes of spectral light. The hole was steadily illuminated with white light from source B (Fig. 1), while the intensity was kept constant at 33.4 photons, *i.e.* the same as the intensity of the surrounding field. The intensity of the spectral light was kept constant at 2332 photons for all wave-lengths except 460 mμ, in which case it was 233.2 photons. The spectral light was interrupted by Disk II, the on-off ratio being 1/44. The results obtained with different spectral stimuli are given in Table II. In general the secondary images following the flashes of spectral light are complementary in color.

The intensity of the light from source B could be gradually decreased to zero to show the transition to the effects described above, which were obtained when only flashes of spectral light were used. Blue and purple secondary images are favored

[5] W. McDougall, The sensations excited by a single momentary stimulation of the eye, *Brit. J. Psychol.*, 1, 1904, 78-113.

by having the intensity of the light coming through the hole from source B lower than that of the surrounding field, yellow and pink by having it higher.

Bidwell observed a phenomenon of the same order with a stimulus pattern like that illustrated in Fig. 2. The disk is divided into black and white halves and has

FIG. 2. BIDWELL'S STIMULUS-PATTERN

an arc whose color may be varied. When the disk is slowly rotated anticlockwise, what Bidwell called momentary negative after-images may be seen following the colored arc. If the lower part of the disk were made white like the upper, the disk would represent the same sort of stimulus pattern as that used in the above experiment. Making one half of the disk black unnecessarily complicates matters.[6]

It can be seen from the data in Table I that, for all parts of the spectrum except yellow, the bright flashes which occur when the rate is 4 per sec. have about the same color as is normally produced by steady stimulation. A yellow stimulus (.580 mμ), however, produces a greenish yellow flash instead of a yellow one. This phenomenon is dependent upon the illumination of the surrounding field, because the flash is seen yellow if the surrounding field is made dark.

Summary. Color sensations obtained with intermittent spectral stimuli have been described. Effects of varying the frequency, wave-length, and purity of the stimulus, and the brightness of the surrounding field have been investigated. At low frequencies bright flashes alternate with dark intervals. The hue of the bright flashes was that characteristic of the wave-length used; the dark intervals were tinged with purple or violet. These purple secondary images are perhaps related to Bidwell's ghost which is obtained with single momentary stimuli and with dark surroundings, although in the present investigation intermittent stimulation and bright surroundings were used and these factors play demonstrable rôles. When the intermittent flashes are admixed to a white area instead of being exposed in a dark area, the secondary images are complementary to the primary ones.[7]

[6] S. Bidwell, On the negative after-images following brief retinal excitation, *Proc. Roy. Soc.*, 61, 1897, 268-271.

[7] Appreciation is expressed to Professor P. W. Cobb for the help he has rendered in directing this investigation.

VISUAL AND KINAESTHETIC JUDGMENTS OF LENGTH

By GERTRUDE RAFFEL, Radcliffe College

During an experiment in which blindfolded *Ss* traced a finger-maze, it was found that a difference in height between two segments of the path felt very large, but looked disproportionately small. Casual observation suggested that most judgments, whether of roughness, length, curvature, or other similar magnitudes, were apt to seem greater to touch than to vision: witness James' observation of the empty tooth socket, which is "believed to be really smaller than the finger-tip which it will not admit, although it may feel larger."[1] In the series of experiments reported below, judgments of length were examined in detail; and it was found that this effect appears consistently and is not disturbed by wide variations in the experimental conditions.

PROCEDURE

O was seated before a table on which stood a screen. He was told that he would be asked to compare the length of two strips of cardboard, and the procedure was explained to him. In a typical series, for example, a piece of cardboard was placed before him, viewed by him, and removed. He then closed his eyes; *E* placed *O*'s right forefinger on the corner of another piece of cardboard; *O* moved his finger along the edge to the left, and compared the two lengths. Equality judgments were avoided whenever possible. The same 4 *Os* were used throughout; they were untrained, and not aware of the purpose of the experiment.

The standard strip of cardboard, except in Series 11, was 1 in. long; the lengths of the comparison strips varied in steps of $\frac{1}{8}$ in. Since these differences are rather large, the threshold could be found readily; in general 6 or 12 judgments sufficed to determine it. When several series were given in the same experimental session, the first was often repeated at the end, and the data in the two instances were always the same. The results may therefore be considered to be reliable.

RESULTS

Series 1. The *Os* felt the comparison strip and then saw the standard one. Although individual differences were pronounced, all the *Os* required a longer strip for vision than for touch.

Series 2. The same stimulus-objects were used as in Series 1, but the order of presentation was reversed: the visual object was presented first. Since the same results appear, the order of presentation is not a determining factor in this experiment.

Series 3. The *Os* looked at both stimulus-objects, and gave accurate judgments.

Series 4. The *Os* closed their eyes and felt both stimulus-objects. The judgments were equally accurate.

Series 5. The *Os* felt the standard strip; then they looked at and felt the comparison strip. The data duplicate those of Series 1 and 2; that is to say, when an object is both felt and seen, as in a situation in ordinary life, the tactual-kinaesthetic experience of length is subordinated to the visual one. This dominance of vision has already

* Accepted for publication October 15, 1934.

[1] W. James, *The Principles of Psychology*, 2, 1890, 178.

been observed; Gibson describes its occurrence when his *O*s wore prisms which made vertical lines bend into curves.[2] A meter-stick held erect seemed shaped like an arc under these conditions; if the *O* looked at the stick while he ran his finger along its surface, it felt curved, even though he knew it should feel straight. This dominance of vision explains why the discrepancy between tactual-kinaesthetic and visual judgments is not more generally and explicitly recognized: when both can be present simultaneously, only the visual experience is perceived.

Series 6. The *O*s saw and felt the standard strip; then they saw the comparison one. Except for one case, vision again was dominant. The results are the same as in Series 3, where the *O*s saw both stimulus-objects without feeling either. Even for the exceptional *O*, the usual difference between touch and vision was markedly reduced.

Series 7. Another series of experiments was undertaken a month later, with the same *O*s, in an attempt to determine the essential factors in the experimental situation. Series 7 duplicates Series 1, and gave substantially the same results. It shows that the threshold of an *O* may increase or decrease during such a long period, even though it is constant throughout one experimental session.

Series 8. Feeling an edge requires more voluntary activity than looking at one; and the subjective enlargement of the felt object might be due to this increase in activity. Such a possibility was controlled, in this series and the next, by having *E* move *O*'s finger along the cardboard for presentation of the standard length, while the comparison length was presented visually. The *O*s were thus equally passive in perceiving both extents, but their judgments showed the usual difference between tactual and visual perception.

Series 9. The reason that a 1-in. stimulus-object always seemed shorter to vision than to touch may be that it could be seen more quickly than it could be felt: an edge felt slowly might show a greater difference from the visual perception than one felt faster. In Series 8, *E* had moved *O*'s finger slowly and at a uniform rate; in this series she moved it as rapidly as possible, decreasing the time out of all proportion to the slight variation which appears in the *O*s' reports. Although the movement was much faster than that which any of the *O*s had used in the control series (Series 7), the average threshold is the same in both. The range of limens in Series 8 and 9 falls within the limits of daily variation of the *O*s, but the difference between them lies in the direction that would be predicted if rate of movement were a determining factor in judgments of length. It may be, therefore, that the rate of presentation has a slight effect, but it cannot be considered the sole cause of the differences between visual and tactual-kinaesthetic judgments.

Series 10. Tactual cues resulting from contact of the finger with the stimulus-object can readily be eliminated. In this series *O* held a piece of cardboard, which *E* placed on the corner of the standard. *O* moved it along the edge of the card and thus formed an estimate of its length without having touched it. The comparison was presented visually. The results show no disturbance, therefore, the essential modality in the tactual-kinaesthetic experience must be kinaesthesis.

Series 11. It is doubtful whether eye-movements occurred in judgments of the rather short lengths used up to this point; the *O*s' estimates were probably based on kinaesthetic cues for one length, and on visual for the other. In this series the

[2] J. J. Gibson, Adaptation, after-effect and contrast in the perception of curved lines, *J. Exper. Psychol.*, 16, 1933, 1-31.

standard stimulus-object was 12 in. long, and the comparison differed in steps of 1 in. All the *Os* required eye-movements to judge the length of the visual object, so that the comparisons were made on the basis of kinaesthetic and kinaesthetic-visual cues. Since vision is the dominant sense, eye-movements accompanying vision would not

TABLE I

TACTUAL LENGTHS EQUATED TO A ONE-INCH VISUAL LENGTH

Since the visual standard is 1.0 the comparison values are ratios. Series 11 is an exception, since the standard was tactual and 12 in. long; but the ratios are also given for this series.

Series	Stimuli	Observers				Mean
		A	B	C	D	
1	(1) touch (2) vision	2.0	1.1	1.2	1.5	1.45
2	(1) vision (2) touch	2.0	1.1	1.2	1.5	1.45
3	(1) vision (2) vision	1.0	1.0	1.0	1.0	1.0
4	(1) touch (2) touch	1.0	1.0	1.0	1.0	1.0
5	(1) touch (2) vision and touch	2.0	1.2	1.2	1.3	1.4
6	(1) vision and touch (2) vision	1.4	1.0	1.0	1.0	1.1
7	(1) touch (2) vision	1.8	1.4	1.2	1.2	1.4
8	(1) slow touch, controlled by *E* (2) vision	1.8	1.4	1.5	1.2	1.5
9	(1) rapid touch, controlled by *E* (2) vision	1.6	1.2	1.2	1.5	1.4
10	(1) mediate touch (2) vision	1.8	1.4	1.3	1.2	1.4
11	(1) touch, 12-in. stimulus (2) vision	26.0	14.5	14.5	21.0	19.0
	touch: vision ratio	2.2	1.2	1.2	1.75	1.6

be expected to affect the judgments; and the data consistently show no diminution in the effect. The slight increase in the ratio of standard to comparison stimulus may be a chance variation, or may indicate that it is a function of stimulus size. This series shows that the important factor is not the absence of kinaesthesis in perceiving the comparison stimulus, but rather the presence of vision.

DISCUSSION

Having shown that a difference exists between visual and tactual-kinaesthetic estimates of length, it is desirable to try to analyze the situation and to determine the crucial factors in it. Time of presentation is apparently unimportant. The tactual-

kinaesthetic experience seems, then, to differ from the visual in (1) that it is tactual, (2) that it is kinaesthetic, and (3) that it requires more activity on the part of *O*. Series 8 and 9 controlled the third of these factors; in Series 10 there were no tactual cues; in Series 11 kinaesthesis was involved in the perception of both objects, but still the customary difference between them appears. Therefore the only essential condition seems to be that vision be present in judging one but not both of the objects. This conclusion suggests that visual impulses may be organized spatially in one projection center, while kinaesthetic impulses are organized in another. Both the visual and kinaesthetic systems, according to such an hypothesis, afford a frame of reference which is consistent within itself, although the two are inconsistent with each other. It is difficult to understand how this lack of correspondence could be due to the training of the *O*, so that the results of the experiment would seem to support nativistic theories of space perception.

APPARATUS

A NEW INSTRUMENT FOR MEASURING PITCH DISCRIMINATION

By RUTH F. WYATT, Northwestern University

The construction of a new instrument for measuring pitch discrimination was motivated by the conviction that the Seashore Sense of Pitch Test and the Kwalwasser-Dykema Pitch Discrimination Test, both of which are given by means of phonograph records,[1] are inadequate for precise measurement of this capacity.

Fluctuation in turn-table speed, not only at the time of recording, but in record reproduction, probably constitutes the most serious source of error, although other questionable features, such as unsuitable selection of increments,[2] variations in intensity,[3] and distraction due to the scratch of the needle and to "metallic overtones"[4] have been indicated in critical literature pertaining to the Seashore Sense of Pitch Test.

Shortly after the appearance of the Kwalwasser-Dykema Music Tests, the physical exactness of the frequencies employed in the pitch discrimination test was questioned by Tilson, who said: "It might be stated as an opinion, supported only by expert judgment of members of the group tested, that the pitch test of the Kwalwasser-Dykema battery is mechanically imperfect. The students in the class tested who were known to have the keenest ears for pitch almost invariably decided that the pitch of the tone was changed on the tests that were supposed to consist of a tone held at the same pitch throughout. Some of these students were mature people who have already made success as solo singers and violinists."[5] Farnsworth, too,

[1] The Sense of Pitch Test is one of a battery of six phonograph records which comprise the Seashore Measures of Musical Talent. They were recorded (five in 1919, the rhythm test later) by the Columbia Phonograph Company and are now distributed through C. H. Stoelting Company, Chicago, Illinois. The Kwalwasser-Dykema Pitch Discrimination Test was recorded in 1930 by R. C. A. Victor as one of ten "Music Tests." To the writer's knowledge, these are the only standardized tests of pitch discrimination which are generally available.

[2] After statistical analysis, R. C. Larson says, "Although the pitch test as it is now, gives us a selection within limits of pitch discrimination, the writer believes that it would be advantageous to get a still finer measurement, especially for more select groups. It is suggested that a new group test be constructed using more increments ranging from 1 to about 8 d.v., and omitting the 0.5 d.v. unit, since it was found to function little if any in a group test. This new arrangement of increments would give us a finer measure of pitch discrimination and would tend to meet more accurately the threshold of a larger number of those tested. An arrangement of increments on this order would very likely eliminate the present skewness of the distribution" (Studies on Seashore's 'Measures of musical talent,' *Univ. Iowa Stud.*, Ser. Aims & Prog. Res., 2, 1930, 53).

[3] P. R. Farnsworth, An historical, critical and experimental study of the Seashore-Kwalwasser test battery, *Genet. Psychol. Monog.*, 9, 1931, 380.

[4] A. R. Gilliland and C. R. Jensen, The reliability of the Seashore phonograph record for the measurement of pitch discrimination, *J. Exper. Psychol.*, 5, 1922, 214.

[5] L. M. Tilson, A study of the predictive value of music talent tests for teacher-training purposes, *Teach. Coll. J.*, 3, 1931, 115.

concluded a study of the Kwalwasser-Dykema Music Tests by stating that several of his ablest subjects (with the best auditory limens) have claimed that the test of pitch discrimination is very inaccurate.[6]

These empirical criticisms seem to be confirmed in an unpublished study of the Kwalwasser-Dykema Pitch Discrimination Test made in May, 1932, by the Acoustic Laboratory of Chicago at the writer's instigation. Oscillograms reveal that in items 2, 4, and 12, in which a tone is supposed to remain constant in pitch throughout its duration, there are discrepancies of from one to seven cycles per second.

In addition to the inaccuracy of the test, it has been criticized for extremely low reliability, low validity, for excessive intensity of the tones and for the rough peremptory voice of the announcer, which has been found disagreeable and distracting. Concluding an intensive study of the Music Tests, Whitley says, "For the pitch test a new recording is greatly to be desired."[7]

The Automatic Tone-Variator

The Automatic Tone-Variator[8] was originally designed to meet a specific need for an accurate test of pitch discrimination sufficiently difficult to differentiate adults majoring in music at Northwestern University. Although up to this time it has been employed primarily in testing the pitch discrimination of groups of adults in various stages of musical advancement,[9] consideration has been given to the possibility of extending its uses to include tests of pitch discrimination on younger *Ss* as well as intensity and time discrimination tests. Incidentally, the Tone-Variator has been very effectively used in demonstrating beats.

The maintenance of uniformity in timbre, intensity, and duration is of great importance in any accurate measure of pitch discrimination. When these accessory factors are dependent upon subjective control by the experimenter, however, discrepancies must inevitably occur. In the Automatic Tone-Variator the frequencies employed, their sequence, intensity, timbre, duration—all these factors—are kept under exact mechanical control and the instrument can be automatically operated. Except for the fact that its tone elements are steel tuning bars, in its operation the Tone-Variator is not unlike a vacuum operated player piano.

Fig. 1 is a photograph of the front section of the instrument with the concealing partitions removed showing the bar assembly and the pneumatic system.

In the upper section may be seen 14 steel tuning bars mounted on resonators and equipped with felt-tipped rubber mallets, dampers, vacuum actions and electromagnets. Steel tuning bars were chosen because, in addition to presenting an easy problem in the matter of mounting and excitation, they can withstand ordinary mechanical usage without derangement of pitch.[10]

---

[6] P. R. Farnsworth, Studies in the Psychology of tone and music, *Genet. Psychol. Monog.*, 15, 1934, 84.

[7] M. T. Whitley, A comparison of the Seashore and Kwalwasser-Dykema music tests, *Teach. Coll. Rec.*, 33, 1932, 749.

[8] Constructed by J. C. Deagan Inc., Chicago, Illinois.

[9] A report of experimental results of pitch discrimination tests made with the Automatic Tone-Variator on professional musicians, advanced amateurs and non-musical adults is to be published soon.

[10] P. T. Young has commented upon the superiority of the bar over the fork for measuring pitch discrimination. Tunable bars and some demonstrations with a simple bar and a stethoscope, *Psychol. Bull.*, 15, 1918, 293-300.

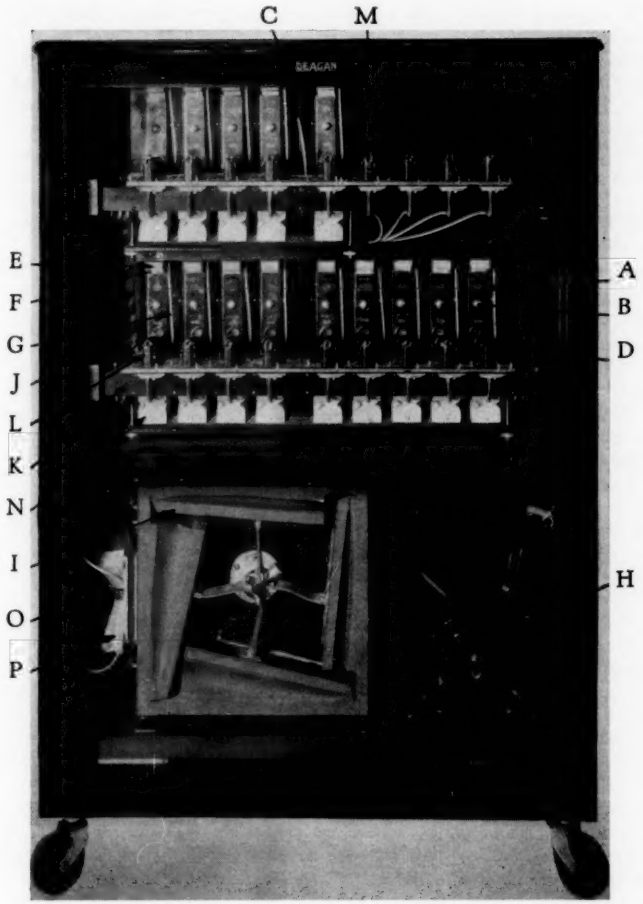


FIG. 1. THE FRONT SECTION OF THE AUTOMATIC TONE-VARIATOR WITH CONCEALING PARTITIONS REMOVED

A, steel tuning bar; B, rubber-insulated pin; C, resonator; D, mallet; E, vacuum damper or bellows; F, damper felt; G, damper vacuum line; H, motor; I, quadruple vacuum pump; J, vacuum line; K, vacuum action rail or header; L, electro-magnet for tonal stimulus; M, electro-magnet for intensity test; N, vacuum action for mallet; O, vacuum pressure regulator valve; P, pneumatic.

A frequency of 465.2 d.v. was arbitrarily chosen as the standard tone. The thirteen comparison bars are higher by 0.5, 1, 1.5, 2, 2.5, 3, 4, 5, 6, 7, 9, 13, 18 d.v. respectively. The first six increments are placed at intervals of 0.5 d.v. while the next four are placed at intervals of 1 d.v. The increments were chosen after considerable experimentation, but obviously others could be substituted or added.

The bars (Fig. 1, A) are anchored at their nodal points (points at which no vibration occurs and determined individually for each bar) by rubber-insulated pins (B). Motion toward the mallet is prevented by felt washers and motion backward is restrained by a linen cord under tension. Four resonators are provided, one for the five bars tuned to standard, +0.5, 1, 1.5, 2 d.v.; another for the bars tuned +2.5, 3, 4, 5, d.v.; the third for the next four bars, +6, 7, 9, 13 d.v., while the +18 d.v. bar, which heretofore has been used only for demonstration purposes, possesses its own resonator (C). All resonators are provided with movable adjustments.

The bars are struck in their exact center by half-hard rubber mallets tipped with Spanish felt (D). Dampers are provided at the top of each bar. They operate in synchronism with the mallet action, being connected by means of the vacuum line (G). The small bellows (E) lifts the damper (F) at the instant the mallet strikes and it releases the damper when the action returns to normal.

To reduce noisiness to a minimum, the motor (H) is specially mounted on rubber blocks in its own frame and the base is mounted on a sheet of sponge rubber. The vacuum pump (I) is quadruple, acting for smoothness of operation. From the main chamber of the pump run numerous vacuum lines (J) to the vacuum action rails or headers (K) which carry the electro-magnets (L and M) and serve two banks of vacuum actions (N).

The spacing of a valve (O) permits regulation of the vacuum within a range of 15 to 60 inches of water gauge vacuum. The pneumatic action not only insures uniformity in the force of the mallet blow, but permits the experimenter to give tests of pitch discrimination at various levels of intensity according to his purpose.

The back of the Tone-Variator, shown in Fig. 2, contains an electric player device, a small 110 volt motor to which it is geared, a hand adjustment for the vacuum pressure regulator valve, a vacuum pressure gauge, a supply of perforated rolls and two keyboards for manual operation of the instrument.

The mallets shown in Fig. 1 are activated in either one of two ways. (1) By means of a key contact (A in Fig. 2 shows the keyboard with one key for each frequency) which enables the experimenter to play the instrument manually. Depressing a key closes an electrical circuit. This energizes an electro-vacuum magnet which opens a valve allowing the vacuum to actuate the mallet to strike the bar, simultaneously releasing the damper. The current available at the key contacts is supplied from a transformer rectifier (not shown) which furnishes 12 volts D.C. for the operation of the magnets. Duration of the tone is manually controlled, the mallets keeping contact as long as the operator's finger depresses the key. (2) by means of the electric player (B, Fig. 2) and a perforated music roll (E) which enable the experimenter to operate the instrument automatically. The perforated roll is placed over the contact roller (D) and as it revolves,[11] the perforations permit

[11] The electric player rotates at the rate of 4.5 r.p.m. but this rate can be varied if desired.

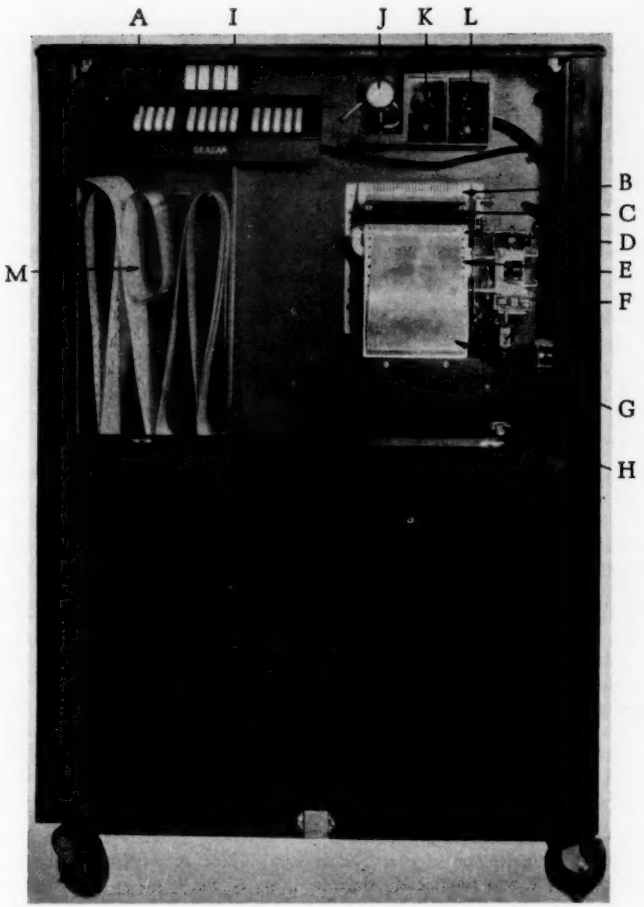


FIG. 2. THE BACK OF THE AUTOMATIC TONE-VARIATOR
A, manual keyboard for tonal stimuli; B, electric player device; C, wire contacts; D, contact roller; E, perforated roll; F, player motor; G, metal roll guard; H, hand adjustment for vacuum pressure regulator valve; I, manual keyboard for intensity test; J, water vacuum gauge; K, pilot light; L, switch; M, extra perforated music rolls in storage compartment.

small wire contacts (C) to touch the roller. This closes the approximate circuits in a manner similar to that described above for the manually operated keys. The experimenter determines what tonal stimuli he will use, the order of their presentation, their duration, and the time interval between them and between trials. A perforated roll is then cut. The lateral position of the slot determines the identity of the bar to be struck, the length of the slot determines the duration of the tone and the distance between slots governs the time-interval between tones. A single trial may consist of two tones as in the Seashore method or of three tones as in the Kwalwasser-Dykema method. Stop slots may be inserted in the roll wherever desired and the entire mechanism can thus be automatically brought to a stop. Tests of 100 trials and 10 practice trials (220 tonal stimuli) have already been perforated on one roll.

The manual method, involving the keyboard, is used for giving preliminary verbal practice in pitch discrimination. The automatic method, involving the perforated roll, is used for the written exercise and for the actual test of pitch discrimination.

By regulating the hand adjustment for the vacuum pressure regulator valve (H), tests of pitch discrimination may be given at any intensity between 15 and 60 in. of water gauge vacuum, the amount of the vacuum being quickly and accurately ascertainable by means of the water vacuum gauge (J).

Operation in measuring intensity discrimination. Only one frequency—the +18 d.v. bar—is used for intensity-tests, but it is struck with five different degrees of force. The four pneumatics (P, Fig. 1) shown at the left of the pump, provide four additional fixed intensities (increments over the standard 20 in. of water gauge vacuum) through varying the closure of the vacuum pressure regulator valve (O, Fig. 1). These pneumatics and their respective electro-magnets (M, Fig. 1) are controlled by their respective keys (I, Fig. 2) and wire contacts (C, Fig. 2). Unlike the keys for the pitch-test, the intensity-keys are similar to organ stops in that they remain depressed until reversed. If more than one intensity-key is depressed, only the greatest intensity is effective. The intensity-keyboard is provided for demonstration purposes but the actual intensity discrimination-test is given automatically by means of perforated rolls. The lateral position of the slot on the perforated roll now determines the intensity of the blow (a standard intensity of 20 in. of water gauge vacuum or one of the four increments) while the length of the slot determines the duration of the tone.

Operation in demonstrating beats. One of the most interesting features of the device is its demonstration of beats, either manually or automatically. In manual demonstration the operator merely depresses two of the keys on the pitch keyboard (A, Fig. 1) simultaneously. If automatic operation is desired, a perforated roll is cut with two parallel slots, causing two mallets to be simultaneously actuated.

Procedure in testing or measuring pitch discrimination. An account of the actual procedure followed in using the Automatic Tone-Variator for group testing of pitch discrimination will serve to illustrate its operation.

(1) The instrument is connected to a 100 volt A.C. outlet and the switch turned on. The hand adjustment is turned until 20 in. of water gauge vacuum is attained. The perforated roll is placed over the roller and the wire contacts are locked down upon it. The Tone-Variator is now in readiness.

(2) A short period is devoted to preliminary practice, the experimenter using the

keyboard. Two tones are sounded successively according to a pre-arranged plan and the *Ss* are asked to judge whether the second tone is higher or lower than the first. Twenty preliminary trials are given by this method and the responses are recited by the group.

(3) Standard blanks are provided for the written exercise and for the test proper. The *Ss* are instructed to check "higher" and "lower." The starting button is pressed and a written practice exercise of 10 trials is played automatically, at the conclusion of which the entire mechanism is stopped by the appropriate slot in the roll. A rest period is allowed during which the *Ss* are permitted to ask questions.

(4) The starting button is pressed and the test proper is automatically administered. The test consists of 100 trials with rests after each 20 trials.

The arrangement of increments in one of the pitch discrimination tests for adults is as follows:

10	trials	employing a difference of	13	d.v.
10	trials	employing a difference of	7	d.v.
10	trials	employing a difference of	5	d.v.
10	trials	employing a difference of	3	d.v.
10	trials	employing a difference of	1.5	d.v.
10	trials	employing a difference of	0.5	d.v.
10	trials	employing a difference of	1	d.v.
10	trials	employing a difference of	2	d.v.
10	trials	employing a difference of	2.5	d.v.
10	trials	employing a difference of	4	d.v.

Each tone is sustained for 1 sec. before it is damped. A time-interval of 1 sec. elapses between the two tones of each trial, and there is an interval of 3 sec. for the response.[12] The test appears to be most satisfactory at a level of 20 in. of water gauge vacuum.

[12] The perforated roll supplies a convenient and exact means for experimental determination of the optimal time relations and of the most desirable arrangement of items.

NOTES AND DISCUSSIONS

Pseudo-Mathematics in the Mental and Social Sciences *

The admirers of a certain very famous psychologist often credit him with having said: "If anything exists, it exists in some degree, and therefore can be measured." "Moreover," say these disciples, "it should be measured. No fact of observation can become a scientific fact until it be measured, and its measure expressed by a number. Only quantitative facts are scientific facts, and we intend to be scientific."

So they say; and so they have set out to measure many kinds of *personal traits*, which they call 'intelligence,' 'emotionality,' 'self-reliance,' 'dominance,' 'introversion,' etc., as well as persistent *social habits* of the individual, which they call 'aptitudes,' 'opinions,' 'attitudes,' 'systematic prejudices,' 'fears,' 'superstitions,' etc. The list of these so-called measurables is long, and it becomes longer whenever a pychological journal issues a new number.

Unfortunately, many psychologists dislike to examine and analyze the fact-finding methods which they employ, fearing, perhaps, that if they should do so, then some one might suspect them of being logicians and philosophers. Hence, some of them have published many studies in a form which they believe to be mathematical, and therefore scientific. Many of their colleagues, being lazy, preoccupied in their own work, fearful of authority, or desirous of peace, have allowed this kind of activity to go unprotested until its results became a nuisance. It is now time to examine them.

Given a set of observational data, it is anti-scientific to subject them to illicit mathematical treatment, just as it would be non-scientific not to apply mathematical treatment if the latter is appropriate. But some kinds of psychological, educational and social data lend themselves to mathematical expression, while other kinds do not. Hence we should always begin by determining what kind of data we have. Thus we must apply a set of universal rules, which belong to elementary logic and elementary algebra, and make sure whether the mathematical treatment that is proposed satisfies *all* these rules or not. If it fails *any* of them it is invalid. Fortunately, these rules are very few and also very simple.

If a collection of observational facts is to be treated mathematically, it is *necessary*, though not *sufficient*, that each fact can be perfectly denoted by a number. But three kinds of numbers lend themselves to denoting, and only one kind to counting: hence, before we perform any operations on the fact-denoting numbers we should determine their kind.

Nominal numbers. We use these numbers each to denote a particular member of the collection, and thus as substitutes for ordinary *names*, as of football players, factory employees, prisoners, railroad cars. If two nominal numbers stand in some particular relation to each other, the fact does not imply that the objects which they denote stand in any corresponding relation to each other. Thus, as Cohen and Nagel remark,[1] from the fact that one prisoner is called "500" and another prisoner

* Read at the St. Louis meetings of the A.A.A.S., before Section I, on December 30, 1935.

[1] M. R. Cohen and E. Nagel, *Introduction to Logic and Scientific Method*, 1934, 294 f.

"100" one could not infer that the first prisoner is "five times as dangerous or wicked" as the second prisoner. Neither could one infer that the first prisoner was sentenced to a longer term than the second prisoner, or even that "500" entered the prison later than "100," since if a convict leaves the prison his number can be reassigned to a newcomer.

Sometimes nominal numbers denote sentences. Thus, in the flag-code of a merchant marine the number 373 might denote, "Yellow fever on board; keep off." By convention, an instructor might write on a student's theme the number 95 to denote, "If you hand in many other papers like this one, I shall recommend you for honors;" or the number 55 to denote, "Unless you go to work effectively I shall exclude you from a continuation-course;" or the number 19 to denote, "You ought to be ashamed to hand in this paper for criticism." Certainly this code does not imply that the first paper exhibited 19/11 times as much information as the second and 5 times as much information as the third; any other set of three numbers could be made to denote the same three sentences as these three numbers respectively denote, if only the code were agreed to. Obviously, no meaning attends the result of any operation performed on nominal numbers, such, for example, as taking their mean.

Ordinal numbers. These numbers, such as "the first," "the second," . . . "the *n*th," are assigned to individual objects, not to denote them as individuals, but to denote the *places* which they severally occupy in some *ordered series.* Note that if two objects in the ordered series should exchange places, they would have to exchange ordinal numbers. To arrange the objects in a collection into an ordered series, one first chooses some one property which all the objects have in common. Thus, Mohs chose the property of *hardness,* which is shared by all the members of his collection of minerals. In general, call the common property *x*. Next, one has to choose some *operation* which defines the relation "*x*-er than." Thus, to define the relation 'harder than,' Mohs chose the operation of *scratching.* If any mineral *A* scratches another mineral *B,* and if also *B* does not scratch *A,* then by definition *A* is 'harder than' *B.* Thus by the defining operation the relation 'harder than' is *asymmetrical.* Moreover, experiment is required to show that the relation "*x*-er than" is *transitive: i.e.* if $A > B$ and $B > C$, then $A > C$. (Here the symbol $>$ means "*x*-er than.") Thus, Mohs showed that if *A* scratches *B* and *B* scratches *C,* then *A* scratches *C.*

If any relation is asymmetrical and also transitive, it permits the objects which it connects to be arranged in an ordered series, otherwise it does not. Thus, Mohs arranged his minerals in the order of their hardnesses, placing diamond, which scratches all the others, at one end, and talc, which is scratched by all the others, at the other end. To diamond he assigned the number 10, to sapphire 9, to topaz 8, . . . and to talc the number 1, assigning to each mineral a number larger than the number of any mineral that it scratches, and smaller than the number of any mineral that scratches it. Thus, carborundum, which scratches sapphire, but which is scratched by diamond, would receive a number greater than 9 but less than 10. These numbers are ordinal numbers: they denote nothing except the *places* in which the several minerals stand in a series that is ordered with respect to the relation *harder than.*

Note that Mohs chose these numbers *arbitrarily.* If he had assigned to diamond

the number 3 and to talc the number 1; or if he had assigned to diamond the number 10 and to talc the number 7, he could still have denoted perfectly the place that each mineral occupies in the hardness-series, simply by following the general rule that we mentioned above. Hence it is meaningless to say that diamond is 10 times as hard as talc, or three times as hard as talc, or 10/7 times as hard as talc, or that its hardness bears any other ratio to the hardness of talc. The relation *harder than*, which is identical with *scratches*, is perfectly denoted by *any* of an infinite number of series of ordinal numbers; and operations on ordinal numbers give meaningless results.

Similarly, if we were to subject say 100,000 persons to the so-called National Intelligence-Test, count the conventional answers that each individual returns, and then ascertain that 70% of the population answered fewer of these questions conventionally than John Walker answered conventionally, and that 35% of the same population answered fewer of these questions conventionally than did William Carpenter, then we might assign to Walker the number 70 and to Carpenter the number 35 to denote their respective centile ratings in the population of tested individuals. But then these numbers would be *ordinal* numbers, and it would be absurd to say that because Walker's place in the series is denoted by a number which is twice as great as the number which denotes Carpenter's place, therefore Walker has twice as much National Test Intelligence as Carpenter has, or that the relations between these two ordinal numbers denote any relation between the 'amounts of' National Test Intelligence which these individuals may be thought of as possessing.

In general, if a property can be *adequately* denoted by nominal numbers or by ordinal numbers, then it is non-additive and non-measurable. Among such properties are hardness, shape, structure, generosity, dominance, radicalness, intelligence and the like.

Cardinal numbers. These and only these numbers express the result of *counting*, and they express nothing except that result. In counting the objects in a collection, one treats each object as if it were interchangeable with every other. Thus one disregards their individualities, which nominal numbers denote, and also their several places in the collection, which ordinal numbers denote. The cardinal number of any collection denotes *how many* members it contains; if its members are *units* of some *distributable* property, then its cardinal number denotes *how much* of the property is distributed among the objects that are considered.

Consider the property called *weight* or *heaviness*.[2] Given a suitable scale-balance, we place in one of its pans a pile of sand and in the other pan two bodies B_1 and B_2 successively. If we should find that B_1 causes its pan to sink while B_2 does not, we assert that B_1 is 'heavier than' B_2. If we should find that the same pile of sand exactly balances another body B_3 and also still another body B_4 when B_4 replaces B_3, we assert that B_4 and B_3 are 'equally heavy.' If we should put both B_4 and B_3 into one pan and add sand to the pile in the other pan until it counterbalances the first, we say that the second pile is 'twice as heavy as' the original pile. Thus, we can define a set of standard weights, and by means of this procedure we can say that

[2] Cf. Cohen and Nagel, *op. cit.*, 289-301.

any member of the set bears some specified relation to any other member. Of course, we are limited by the capacity of the instrument and also by its construction. We feel sure that if these limitations could be removed, or if we could replace this instrument by other instruments in certain ranges of weight, or if we could correct for friction, distortion, etc., then these principles would hold perfectly both within and without our range of experience. But this is pure assumption.

These operations suggest that heaviness differs from hardness in being an *additive* property. Nevertheless, we must not be too hasty in assigning numbers to two or more bodies according to their operationally defined weights, and in assuming that the result of adding these numbers expresses the result of placing these bodies in the same pan of the balance. In general, if a property is truly additive, and measurable, it must satisfy all the following criteria.

(1) The relation 'x-er than' must be unequivocally defined by the behavior of the detector of the property x, and in such wise that this operational definition shows the relation to be *asymmetrical*. Given a collection of objects which have the property x, the detector must show that between any two objects in the collection, such as B_i and B_j, one and only one of these relations holds: (a) $B_i > B_j$; (b) $B_i = B_j$; (c) $B_i < B_j$; with respect to x-ness. Thus, if x is weight, we assert (a), (b), or (c) according as B_i counterbalances more sand than B_j, the same sand as B_j, or less sand than B_j counterbalances, when B_i and B_j are interchanged in the detector. Outside a determinable range of uncertainty these judgments are not confused; if this range is greater than is permitted by the degree of precision which we demand, then the operation of counterbalancing on this instrument does not define the relation 'heavier than.' But if the detector permits the unequivocal judgment $B_i > B_j$, then it always yields the judgment $B_i \neq B_j$ and $B_j \not> B_i$, and thence the judgment $B_j < B_i$. Thus the relation 'heavier than' is *asymmetrical* if it is *determinate*.

(2) The defining operations must show that the relation 'x-er than' is *transitive: e.g.* if $B_i > B_j$ and $B_j > B_k$, then $B_i > B_k$.

(3) The collection of objects which have the property x must in its turn have the so-called *group-property*.[3] For example, if two objects B_i and B_j severally excite the detector of the property x, then the collection must contain another member B_k, such that $B_k > B_i$ and $B_k > B_j$ and $B_k = B_i + B_j$ in the sense that B_k excites the detector of the property x in exactly the same manner as B_i and B_j together excite it. In other words, if the effects of B_i and B_j on the detector are additive, then the detector must show an effect which is equivalent to their sum. This implies that the number of objects which have the property is *infinite*.[4]

(4) The addition in the detector of bodies that have the property x must satisfy the *commutative law* of the addition of numbers. Thus, if $B_i + B_j = B_k$, then $B_j + B_i = B_k$.

---

[3] Cf. C. J. Keyser, *Mathematical Philosophy*, 1922, 202.

[4] What about adding *negative quantities* of the same property? Strictly speaking, no such quantities exist. A surface, for example, cannot carry fewer units of positive charge than none. When we speak as if it does, we mean that we must add some positive charges to the surface in order to bring the detector into a certain equilibrium state. We interpret this as implying that the surface carries more negative charges than positive charges; all that we detect is the excess of the one kind of charge over the other. We cannot detect the absolute quantities of either kind of charge. The same principle applies to any other instance that we may select.

(5) The operation of empirical addition must satisfy the *associative law* of addition of numbers. Thus, $(B_1 + B_3) + B_k = B_1 + (B_3 + B_k)$, etc.

(6) The operation of empirical addition must satisfy the *axiom of equals* which holds for the addition of numbers. Thus, if $B_1 = B_1'$ and $B_3 = B_3'$, then $B_1 + B_3 = B_1' + B_3'$ and $B_1 + B_3' = B_1' + B_3$.

In particular, if the property x has an antagonist $\bar{x}$, which we agree to call the negative of x, then the operation of empirical addition must satisfy two additional criteria; namely:

(7) The collection of objects, ordered with respect to the relation 'x-er than,' must contain a member B_0 which does not excite the detector of the property x, and which, if placed in the detector along with another object B_1 will always yield the result $B_0 + B_1 = B_1 + B_0 = B_1$. Hence B_0 is called the *neutral member* of the collection.

(8) Finally, if the collection contains an object B_1 which excites the detector of the property x in one manner, it must contain a corresponding object B_{-1}, which excites the detector in the opposite manner, and such that if B_1 and B_{-1} are placed together in the detector, they do not excite it. Since the sum of their several effects on the detector is imperceptible, as is also the effect of B_0 upon it, we say that $B_1 + B_{-1} = B_0$.

If a property is measurable it satisfies all these criteria; if it fails any of them it is non-additive and non-measurable. In other words, the addition of 'quantities' of the property x has to satisfy all the axioms of addition of cardinal numbers: otherwise it cannot be expressed by the result of numerical addition.

It is at least doubtful whether any observable property exists which *satisfies* all these axioms. For example, the class of all weights that are detectible by means of the scale-balance does not have the group-property *(3)*; and its members do not, in general, obey the commutative law of addition *(4)*; both criteria remain unsatisfied if the balance is overloaded. Nevertheless, there is a range of weights, which we can determine empirically, and within which we can *pretend* that scale-balance weight is additive, and prove that we have not so introduced an uncertainty which exceeds some specified standard of tolerance. Now some mental and social data are of this kind: they *approximate* the criteria of measurability within a given standard, although they do not *satisfy* the criteria. But other kinds of mental and social data fail them utterly. We shall mention examples of both kinds.

Are perceptible brightnesses measurable? Perceptible brightnesses are not identical with what the physicist calls the brightnesses or luminosities of surfaces. To determine the physical brightness of a surface, one first ascertains the wavelengths of the radiation which the surface emits, transmits, or reflects toward the detector. Next, with respect to each minute range of wavelengths, one determines (a) its surface-rate of power, and (b) its so-called 'visibility factor,' taken from a standard wavelength-luminosity curve. Taking the product of (a) and (b) for each minute range of wavelengths, one then summates the products. These operations define the physical brightness of the surface. It is imperfectly correlated with perceptible brightness within certain limits, but it is not identical with the latter.

Consider a surface illuminated by two sources S_1 and S_2 in succession. When S_1 is used alone, the observer perceives a brightness of the surface which we may call B_1; when S_2 is used alone, he perceives a brightness B_2 on the same surface. Using the flicker-method of photometry, or else the method of direct comparison,

let us balance B_1 against a comparison-field B_1', and also balance B_2 against another comparison-field B_2'. Now, expose the surface to both sources S_1 and S_2 at once. If we agree that in so doing, we add B_1 and B_2, then by the axiom of equals *(6)*, $B_1 + B_2 = B_1' + B_2'$, and $B_1 + B_2' = B_1' + B_2$. But this is not generally true. Suppose, for example, that the sources which produced B_1 and B_2 respectively, emitted only lithium light ($\lambda = 671$ mμ), while the sources that produced B_1' and B_2', respectively, emitted only thallium light ($\lambda = 535$ mμ). Suppose moreover that $B_1 = B_1'$ is high, while $B_2 = B_2'$ is low. Then $B_1 + B_2'$ is the sum of a bright red and a dim olive-green, while $B_1' + B_2$ is the sum of a bright olive green and a dim red. Although observation yields the separate equations $B_1 = B_1'$, $B_2 = B_2'$, it is very likely to yield $B_1 + B_2' \neq B_1' + B_2$. It may also yield $B_1 + B_2 \neq B_1' + B_2'$. The operations may not satisfy or even approximate the axiom of equals. Some have suggested that heterochromatic brightnesses *would be* additive *if* it were not for Purkinje's effect, and they may be right: anything probably would be different from what it is if only it did not remain the same. But since Purkinje's effect characterizes the brightnesses that we perceive, anything that lacked it would not be a perceptible brightness.

But are perceptible brightnesses measurable if they are *homochromatic?* No, their addition is not *commutative (4)*. Suppose that direct comparison yields $B_1 > B_0$, $B_2 > B_1$. Suppose also that B_2 is *much* greater than B_1. Then, direct comparison may yield $(B_1 + B_2) > (B_0 + B_2)$; $(B_2 + B_1) = (B_2 + B_0)$; whence, if $(B_0 + B_2) = (B_2 + B_0)$, then $(B_1 + B_2) \neq (B_2 + B_1)$. Some have surmised that the axiom of equals would be satisfied, together with the commutative law, if it were not for the fact that brightness-perception goes by jumps; nevertheless, differential thresholds characterize brightness-perception, and anything that is not infected with them is not perceptible brightness.

Again: has the class of all perceptible brightnesses the so-called *group-property?* If so, then the addition of any perceptible brightness B_i to any other perceptible brightness B_j must be equivalent to another perceptible brightness B_k, such that *(3)* $B_k > B_i$ and $B_k > B_j$. By definition of perceptibility, if B_i and B_j are perceptible while B_o is not, then $(B_i + B_j) > (B_i + B_o)$, and $(B_i + B_j) > (B_o + B_j)$. But if either B_i or B_j is near the 'terminal threshold' of perceptible brightness, direct comparison is very likely to yield that $(B_i + B_j) = (B_i + B_o)$ or that $(B_j + B_i) = (B_j + B_o)$; whence $B_i = B_o$ or $B_j = B_o$, which contradicts the hypothesis.

Hence, perceived brightnesses are not, in general, additive or measurable. Like scale-balance weights they fail some of the criteria of measurability, although, within certain limits, they approximate these criteria within some determinable standards. If we are content with these standards, then we can set limits to a range of brightnesses within which we can pretend that brightnesses are measurable and outside which we cannot pretend that they are. But we now have to consider some classes of mental and social data of which not even this is true.

Are perceptible hues measurable? First, can we arrange all perceptible hues in the order of their resemblance to some standard hue, such as Helmholtz's primary chlorine green? Yes; if we exclude neutral gray from the collection, we can so arrange them, but not in a rectilinear series, as axiom *(1)* requires. The arrangement must correspond to a closed curve, on which the purple that complements the standard chlorine green will be at the pole opposite the latter, since it resembles

the chlorine-green less than any other hue resembles it. But the relation 'resembles' is non-transitive, and so violates axiom *(2)*: for example, one may find a red and a blue green which resemble the chlorine green, but which (being complementaries) do not resemble each other. We have noted that neutral gray is not in this ordered series. But unless it belongs in the collection, axiom *(3)* is not satisfied, for the rule that the addition of one hue to another gives a third hue requires that the neutral hue be in the collection, otherwise the rule does not provide for the addition of complementary hues in certain proportions. Moreover, unlike the series of natural numbers, the number of distinguishable hues is finite. For example, there is a chlorine green which is "chlorine-greener" than any other hue, and there is a purple which is less chlorine-green than any other hue. The same assertion holds in principle for any primary hue that we may select. Axioms *(1)*, *(2)* and *(3)* are violated: *perceptible hues are non-additive and non-measurable.*

Consider next a class of *social facts,* which some authors falsely have treated as measurables: *e.g.* attitudes, interests, intelligences, aptitudes, skills, drives, and the like. Thurstone,[8] for example, has asserted that attitudes can be measured. Let us look into this question.

We say that a person holds one attitude or another toward some specified change (now occurring or in prospect), according as he is predisposed to make one kind of response or another kind of response to the change. Thus an attitude is a preparation for action. We suppose that it is correlated with a corresponding pattern of tensions and conductivities in the nervous system, and thus determines the *equilibrium state* of the latter. Just as solid carbon obeys one set of thermodynamic laws when it is in the form of soot, another set when it is in the form of graphite, and still another set when it is in the form of diamond: so a person, while he holds the 'attitude' of a pacificist toward a war, will make certain responses to banners, posters, pageants, poems, and to appeals to enlist, to buy bonds, to encourage the government, etc., which he would not make if he were holding the attitude of a patriot toward the war. Of course, we cannot describe his attitude physically, probably we would not so describe it even if we could. Rather, we *name it after* the course of action which we think it predisposes the person to execute, but the properties of the attitude do not depend on our way of describing it.

These assertions together imply that there are as many attitudes as there are distinguishable patterns of action, or distinguishable patterns of brain-tension. Each attitude predisposes the subject to behave according to its own corresponding set of laws. It is meaningless to ask *by how much* one attitude differs from another as it is meaningless to ask *by how much* the molecular arrangement of carbon in the form of graphite differs from its arrangement in soot or in diamond.

Thurstone does not make the mistake of supposing that these questions are meaningful, but he does assert that it is possible to measure some of the properties of attitudes, such as their *predisposing tendency* toward a given course of action, such as actively participating in a war. He assumes, moreover, that this tendency is additive, in the sense that by taking the sum, or else the mean, of the tendencies of two or more attitudes toward a given course of action, he gets the equivalent of

[8] L. L. Thurstone, A law of comparative judgment, *Psychol. Rev.*, 34, 1927, 273-286; Attitudes can be measured, *Amer. J. Sociol.*, 33, 1928, 529-554; Theory of attitude measurement, *Psychol. Rev.*, 36, 1929, 222-241.

the predisposing tendency of another attitude toward the same course of action. This assumption I invite you to examine.

Before we can proceed to *measure* any property of any attitude, we have to *detect* the attitude itself. But why is it proposed to detect a person's attitude toward a given change, in the first place? Most attitude-testers are practical men: they wish to know, for example, what one needs to do to a person to cause him to execute some course of action, such as joining the church, voting against a bond-issue, buying accident-insurance, enlisting in the army, or the like. Or, if they believe that some tendencies toward action are not amenable to propaganda, then they wish to know what means of persuasion they must avoid. Hence they wish to know what the person is already set or tending to do about the issue; or, in other words, to detect his attitudes. But, unless they intend to start working on him as soon as they have finished the test, if they use the results of the test, they presuppose that he will be holding on another occasion the attitudes which he held when they tested him. Unless the second occasion is dated, they presuppose that the test indicates what attitudes the person most probably will hold on *any* other occasion within a reasonably short time. In other words, they presuppose that the result of a single test will indicate what the person's attitudes are, *habitually*.

This presupposition is not, in general, plausible. The sinner today may become a missionary before tomorrow night; a man who is ready to give all for the love of his lady today may be glad within a fortnight that she quit him. But if the practical man denies the presupposition, he thereby denies that the detection of attitudes is useful to the propagandist.

How do the attitude-testers try to detect what a person is tending to do about a given change at the moment? *They ask him questions!*[6] Does what he *says* indicate what he is tending to *do?* Thurstone says merely that *if* it does, *then* the tester, by taking the average of the numbers which many judges set down to denote 'how favorable' toward the question they guess his statement of opinion to be, can determine 'how strong' is his predisposing tendency with respect to a corresponding *action*. The 'predisposing tendency' is not operationally defined; those who adopt this hypothesis thereby assume that the average of the guesses of the individual judges indicates the result of a set of operations that are either inconceivable, infeasible, or else merely not yet performed. Thurstone explicitly refuses to make this assumption, but his procedure is pointless unless what a person *says*, at one instant, about an hypothetical situation reliably indicates what he is most likely to *do* if the situation should *ever* become actual; this he *implicitly* presupposes.

Thurstone assumes that all possible opinions toward a given question (and with them, the corresponding tendencies to action) can be arranged in a series that is ordered according to the relation *more favorable than;* that 'favorableness' is additive according to the laws of algebraic addition; and moreover, that if B_1 and B_2

[6] I once knew a certain devout Methodist who used to testify that his local church was more precious to him than all his worldly possessions, and who contributed less money toward its annual support than he spent in half a year for chewing tobacco. Also, I know certain individuals who talked pro-German until 1917, pacificistically through 1917, but who entered the army in 1918 and fought, although they were entitled to exemption. And I know certain capitalists who praised the New Deal ardently in 1933, talked communism in 1934, and raised money for the Liberty League in 1936.

are *any* two opinions in the ordered collection, the latter contains another opinion B_3, such that B_3 is more favorable toward this course of action than is B_1 and less favorable than B_2 toward it. This assumption implies that between any two ordered opinions there are an infinite number of similarly ordered opinions, so that the numbers that express their 'favorableness' constitute a 'dense set.'

Consider these: If war should be declared, I would fight,

(A) under all conditions;

(E) under no conditions;

(I) under some conditions; namely, (a), (b), (c), ...;

(O) not under some conditions; namely, (a′), (b′), (c′),.....

Obviously the classes (A) and (E) of opinions contain only one member each; classes (I) and (O) may each contain many opinions. Note especially that (A) and (O) contradict each other, as do (E) and (I). The axiom of excluded middle implies that between the sole opinion in class (A) and the least unfavorable of the opinions in class (O), no opinions can be inserted in the ordered series; and that between the sole opinion in class (E) and the least favorable opinion in class (I) no opinions can be put into the ordered series. *Opinions do not stand in a continuum.* If opinions are the transforms of attitudes, then *attitudes do not stand in a continuum.*

We cannot now review the elaborate procedures which Thurstone employs[7] for ordering statements of opinion respecting any given issue, and for assigning scale-numbers to the statements. The careful reader will discover clearly that *each procedure merely defines a rule for assigning* ordinal *numbers* to the statements. Nothing in the procedure provides a means of determining 'how much' favorableness any statement indicates, or 'how much more' favorable to a given decision one statement of opinion is than another.

From another direction let us examine the assertion that a person who endorses two or more statements expresses the sum of the 'favorablenesses' of each. Is any statement more favorable to the prospective war than the sole statement in (A)? No, for it includes all the statements in (I). Hence the property of 'favorableness to war' is non-additive, because the objects which have it do not form a collection that has the *group property.* If a person endorses (A) and also, for example, $(I.\ c)$, he has not indicated more 'favorableness' than if he endorsed only (A). But if both are indicated by finite scale-numbers, their sum would be greater than the scale-number of (A). If one averages the scale-numbers, as Thurstone recommends because of the possibility of ambiguity in both question and answer, then the individual who endorsed (A) and also every member of (I) separately would earn a lower patriotic score than a person who endorsed only (A). Perhaps the procedure should be worked over.

In brief, no property of any attitude can be measured unless the attitude can be first detected. It is not evident that it can be detected from the person's own assertions of 'opinion.' Nor does the procedure which we mentioned provide a means of measuring the 'favorableness' of opinions themselves. It yields a counterfeit measure of the strength of a person's tendency toward any course of action. Similar defects inhere in the so-called measurement of interests, drives, skills,[8] X-O emotional stability, intelligences, and the like.

---

[7] *Op. cit.*, cf. footnote 5.

[8] H. M. Johnson, Some neglected principles in aptitude testing, this JOURNAL, 47, 1935, 159-165.

But why bother about principles? If these procedures are unsound, surely the 'Movements' will die of themselves! Why not let them alone?

The logician may answer, "Because if an invalid method leads one to factual truth, it is only by accident." The practical man may add, that the time and effort that are wasted in futile endeavor are subtracted from the amount that is available for effective work. There are questions concerning human traits which are interesting; they may be important; they may be answerable; but they probably will remain unanswered as long as they are attacked by attempts to measure what is intrinsically non-measurable.

For example, consider the attempts made during the past 30 years to determine certain psychological effects of drugs, partial asphyxiation, fatigue, insomnia, and sleep. The plan of nearly every experiment presupposed that the effect on some arbitrarily selected function would be *graded* according to the magnitude or the duration of the agent of impairment. In nearly every instance no such gradation was found; whereupon the investigator drew the conclusion that the agent was ineffective. The conclusion was false, because it rested on the false presupposition, that if the agent had an effect, the effect would be of the kind that he was seeking. But if the experimenter, like Dunlap,[9] and Bagby[10] and others in their study of asphyxiation, had asked, not 'how much' the function was affected, but *in what manner* it was performed under the compared conditions, he probably would have had to draw a different and valid conclusion. Elsewhere,[11] I shall present this field of investigation in detail. For the present, it is enough to say that it illustrates, clearly and almost tragically, the practical waste that results from reliance on false presuppositions and on pseudo-mathematics. Those data should be measured which can be measured; those which cannot be measured should be treated otherwise. Much remains to be discovered in scientific methodology about valid treatment and adequate and economic description of non-measurable facts. Their detection as such, however, is logically simple.

American University H. M. JOHNSON

ERRATA

Professor Edmund Jacobson wishes to correct the following errors in his article on "The course of relaxation in muscles of athletes that appeared in the last number of the JOURNAL. On p. 99, line 37, the numbers should be "1/3 or 1/4" in place of "3 or 4." The sentence should then read, "String tension in the galvanometer is so adjusted as to yield an excursion of 1/3 or 1/4 cm. per millivolt applied to the string terminals."

Professor Karl M. Dallenbach also wishes to correct two errors that were made in his apparatus note, "Two new A.C. chronoscopes." On p. 147, the caption to the ninth column of Table II should be "10th 10" instead of "4th 10," and the legend of Fig. 1 on p. 151 should read "Dial of Model SWC-1" instead of "SC-1."

---

[9] Knight Dunlap, Psychological research in aviation, *Science*, 49, 1919, 94-97.
[10] English Bagby, The psychological effects of oxygen deprivation, *J. Comp. Psychol.*, 1, 1921, 97-113.
[11] H. M. Johnson, *Human Sleep*, (in press).

BOOK REVIEWS

The Evolution of Modern Psychology. By RICHARD MÜLLER-FREIENFELS. Translated from the German by W. Béran Wolfe. New Haven, Yale Univ. Press, 1935. Pp. xvi, 513.

While it is doubtless foolish to expect the publisher's jacket to be objective and judicious, there is no reason why it should lead us to expect a totally different kind of book from what it encloses. This book is not "the most complete discussion of modern psychiatry and psychology that has been written." It is hardly a survey of psychiatry at all. In fifty pages it shows the coincidence of development between psychology and its truant offspring, psychiatry. It does not provide "the distracted layman" or "psychologists themselves" with "a wide knowledge of psychological trends in all countries." The author indeed apologizes for his relative neglect of "the vast results of psychological research in the United States" but other countries fare little better. Sixty-one percentage of the references are German (as against 21% and 17% respectively for English and Romance language references) and the discussion centers almost entirely around the development of contemporary German psychology.

This is neither surprising nor regrettable. The recent developments in Germany are so foreign to our own current preoccupations that—with the exception of *Gestaltpsychologie* —they have remained pretty completely misunderstood. It is well to have this survey which puts the development in its historical setting and thus helps to render it intelligible.

We must not, then, expect this book to illuminate our own post-behaviorist situation. Müller-Freienfel's selection is designed for another purpose.

A single item may illustrate. As early as 1889 Bergson was attacking the concept of the intensive threshold as conceived by Delboef. We do not actually measure the intensity of consciousness at all; we merely measure estimates of the intensity of external stimuli. For Müller-Freienfels as for Bergson this points to a limitation of the psychology of pure experience. Sensory intensities are qualitative and consciousness cannot be quantified.

The withers of most contemporary psychophysicists are of course quite unwrung by such strictures. They do not believe themselves to be quantifying consciousness; they are engaged in discovering how well an organism can perform certain tasks —*inter alia* the task of estimating the intensity of objective stimuli. A consideration of Bergson's criticisms may help the current psychophysicist to understand what he is doing. In order to move forward we often need to turn the light of the past on the path we are treading. Müller-Freienfels, however, turns the light of the past not on our path but on the path of a radically divergent psychology.

The value of the book thus lies mainly in the understanding it yields to our puzzling German contemporaries. Regretfully the reviewer must record his judgment that this psychology is not terribly worth understanding. I am not likely to be accused of undervaluing theory. I can agree with the author's dictum that "all experiment is preceded by theory; and the history of all sciences is primarily a history of their theoretical approach to certain problems" (p. xv). Müller-Freienfels, however, himself goes further, and describes German psychology as going further.

Again and again he refers with apparent approval to the creation of a system of psychology by taking a certain philosophy as the point of departure.

Now no one denies that systems of psychology are related to philosophy, even that in a sense they are philosophy. The scientific movement, however, is committed to the view that the system must grow out of the facts. We are not naïve about facts. We are aware that facts are segments carved from reality somewhat according to individual temperament. We insist, none the less, that it be reality not figment which is formed into theories. Unless Müller-Freienfels misleads us, an important faction of German psychology neglects this important postulate, another explicitly denies it. Science is regarded as of value only as the handmaid of intuition. German psychology has gone dialectic.

Even so there might be something in it for us. Dialectics can help to define the problem, though they can never solve one. The reviewer has hoped that German characterology, for example, might be of service in reorienting us. We have spent much energy in the attempt to make ratings reliable without much thought as to what we were rating. A dialectic and intuitional approach might *suggest* the units or categories in which personality or character is to be described. Alas, the Germans also seem more interested in finding what elements of physique or of behavior are significant before they have determined what there is psychologically to be signified. As for the typological schemata of Dilthey, Spranger and company, for all their subtlety in describing *Weltanschauungen* and *Lebensformen,* they remain completely naïve as regards any organic or psychological basis.

The Evolution of Modern Psychology thus dashes the hope of one thoroughgoing eclectic for any fruitful interaction between the 'structural' psychologies of Germany and the psychology of performance. This disappointment colors one's appreciation of a book very well done within its self-determined limits.

Moreover, for special topics there are excellent surveys. The brief account of psychiatry is enlightening. The treatment of the analytic schools is clear. The six chapters on the "Superindividual Psychic Life" are distinctly valuable, though more sociological than psychological.

One must sympathize with certain difficulties of the translator. "Seele" of course is not soul but neither is it any of the other alternatives. One can only say that "soul" affronts us rather more than "mind." "Anschaulich" is certainly seldom merely "visual"; have not "palpable" or "tangible" about the same metaphorical value? "Bewusst" is certainly not usually to be rendered "the conscious." Yet on the whole the translation seems adequate.

Ohio State University HORACE B. ENGLISH

A Dynamic Theory of Personality. By KURT LEWIN. New York, McGraw-Hill Book Co., 1935. Pp. ix, 286.

This book is a collection of eight articles, only two of which have not appeared previously in print. Two of the previously published articles have been available in English heretofore, the remainder having appeared originally in German.

The first essay, "The conflict between Aristotelian and Galileian modes of thought in contemporary psychology," was published originally in the *Journal of General Psychology.* In it, Dr. Lewin develops the thesis that Gestalt psychologists are primarily Galileian in their thinking as contrasted to the Aristotelian flavor of the

mental processes of non-Gestalt psychologists. (By so classifying modes of thought Dr. Lewin is guilty of that very dichotomizing which he ascribes to Aristotelian thinking.)

The second essay, "On the structure of the mind," is an extract from the introduction to a series of articles, edited by Dr. Lewin, published in the *Psychologische Forschung*. In this chapter, the author attacks the association doctrine and considers the psychical sources of energy. The third chapter, "Environmental forces in child behavior and development," is probably the most familiar of Dr. Lewin's writings to American psychologists since it is a reprint of his chapter in the second edition of *The Handbook of Child Psychology*.

The fourth essay, "The psychological situations of reward and punishment," which originally appeared in German is the most interesting and readable of the essays. In it, the author considers the important question of how rewards and punishments can be utilized to induce or suppress behavior. Chapter five, "Education for reality," is a short presentation of Dr. Lewin's viewpoint applied to education and should be of particular interest to progressive educators.

Chapter six, "Substitute activity and substitute value," has not been previously published. A very ingenious, though simple, experimental technique for investigating the substitute value of activities is described. Chapter seven, "A dynamic theory of the feebleminded," contains a discussion of the relation of will and needs to intellectual deficiency and evidence that the feebleminded child is more distractible and less satisfied by substitute activity than the normal child.

Chapter eight, "A survey of the experimental investigations," which ought to give flesh to the theoretical skeleton contrived by Dr. Lewin, is very inadequate. It should have been expanded in order to demonstrate more clearly the factual basis upon which Dr. Lewin's theory rests. The interested reader who possesses a knowledge of German can, of course, refer to the original articles of which this chapter is a summary, but that should not have been necessary.

The attempt to explain behavior in terms of the interaction of intraorganic or psychical factors and environmental forces is a commendable one, if the postulated factors and forces are measurable. The reviewer believes that the principal goal of science is the establishment of functional (dynamic) relationships between variables. If behavior is to be explained in terms of need, tensions, forces, differentiation and connection of energy systems, changes in meaning and the like, then these conditions should be measurable, either directly or indirectly, and independently of the activity to which they are being dynamically related.

A child, for example, spies candy on a nearby table, approaches and eats it. For Dr. Lewin, the child's behavior in this situation is explained by stating that the sight of the candy taps an energy system or need which supplies the force necessary for the resulting activity. Are independent measures of these dynamic factors available or are they entirely inferential? At the present time many of Dr. Lewin's concepts or hypotheses do not seem amenable to experimental verification.

This drawback to Dr. Lewin's system is an error of omission rather than commission. The reviewer does not suppose for a moment that Dr. Lewin is unaware of the unfinished nature of the evidence upon which his theory is based. The experimental work which his ideas have inspired is sufficient indication that measurements of these inferred factors are being made. For otherwise experiment would be impossible.

It is the reviewer's opinion that this book represents a most significant contribution to psychological theory and method. It demonstrates a high order of insight into the dynamics of behavior and is additional proof of the fact that psychology is being vitalized by the organismic viewpoint. The reader may object to Dr. Lewin's frequent use of anecdote to illustrate his points. Anecdote is, however, the original stimulus for much scientific research. The reviewer will be very much surprised if considerable experimental work in relatively unexplored fields is not initiated by Dr. Lewin's theory. The experimental investigations so instigated will demonstrate the fertility if not the correctness of Dr. Lewin's ideas.

In concluding the reviewer congratulates Adams and Zener for the splendid manner in which these essays have been translated. Anyone who has tried to read Lewin in German can appreciate the difficulties which the translators must have encountered. Whatever may be the influence of Lewin's writings on his American audience will be due in large part to the work of these two translators.

University of Oregon CALVIN S. HALL

Children's Sleep. By S. RENSHAW, V. L. MILLER, and D. P. MARQUIS. With an introduction by W. W. Charters. New York, Macmillan Co., 1933. Pp. xix, 242.

The authors studied 79 girls and 92 boys, all being inmates of an institution. Their ages ranged between 5 and 20 yr., through a middle range of 10.3 to 15 yr., and a median of 12.8 yr. Their so-called IQs ranged between 30 and 120, through a middle range of 79 to 99, and a median of 89. The authors sought for a typical picture of the nocturnal "motility" of the sleepers, and of its changes with certain systematic variables, one of which was certain motion-pictures presented in the evening.

First, *they detected each minute during the night in which the sleeper's bodily movements actuated the pen of a kymograph.* The pen was controlled by an electromagnet in the circuit through a stationary brush and a movable commutator-bar attached to the bed-spring so as to be displaced vertically by the sleeper's movements. Each segment of the bar was 3.0 mm. wide, and all the conducting segments were in parallel. In certain critical positions of the brush on the bar, an infinitesimal displacement in one direction would excite the pen, whereas a displacement of 5.5 mm. in the other direction was required. They wished to record all gross changes, and only gross changes, in the sleeper's bodily position, but evidently they recorded some very slight changes also. The extent of travel of the bar does not correspond either to the magnitude or to the frequency of displacement of the pen; hence, the record does not distinguish gross movements from slight movements of the sleeper.

All the beds appear to have been alike, the mattress resting on a system of cables fastened at their ends to the frame through helical springs. Each cable was clamped rigidly at a few points to either of its nearest neighbors, and to no others. The commutator-bar was attached to the central cable, near its middle: hence, a person, in making a given movement, will displace the bar farther if he lies near the center of the bed than if he lies near its edge. Thus, there is a critical class of movements, any of which will displace the pen if the subject occupies any of a certain class of locations in the bed; not if he occupies any location not belonging to this class. Hence, in the long run, the larger and heavier the sleeper

is, the more of his movements will be registered. Moreover, if the class of children between 14½ and 18 yr. of age contains relatively more large and heavy sleepers than the class of 10½- to 14½-yr.-olds, and relatively still more than the class of 5- to 10½-yr.-olds, we should expect the proportion which the number of detected movements bears to the number of actual movements to increase according to the age-range of the several classes of sleepers. Now the authors *found* that the older subjects *registered* movement more frequently than the younger subjects. They *assumed* this finding to mean that the older subjects not only registered, but also made, more movements than the younger. They left this assumption untested. I question its factual truth.

Secondly, the authors counted the active minutes in each hour, and presented the count separately for each hour of the night. If the same subject slept many nights under comparable conditions, they averaged the hourly count night by night; if many subjects in one age-group slept under comparable conditions, they averaged their individual averages. On the latter procedure they base an important interpretation, although on p. 71 they first deny that the procedure can be legitimately applied to their results.

In their most typical experiment (V) they report (p. 134) that on movie-nights, in the average, the boys stirred in 1.24 *more* minutes per hour, and the girls in 0.2 *fewer* minutes per hour, than on normal nights. They present these differences along with determinations of their standard errors, neither standard error being computed correctly from the data in the table. The difference returned by the boys is actually 1.18 (not "3.786") times its own standard error, while the difference returned by the girls is 0.018 (not "0.015") times its own standard error. The authors discuss the difference found for the boys, and also the difference between it and the difference returned by the girls, as if the boys' difference were statistically reliable, whereas it is properly attributable to chance. If it were reliable, it would signify that "the average boy," spending 9 hours a night in bed, stirs in 11 more minutes of that time on movie-nights than on normal nights. What kind of good, and how much of it, would "the average boy" have got out of these 11 minutes if he had spent them in rest? I wonder!

The authors compare these differences with those which *other groups of subjects* return, between their average motility on odd-numbered and even-numbered nights, and between their averages based on many normal nights and their averages based on a few holiday-nights. None of the differences, however, is significant.

In the same experiment it appears from Table 31, p. 125, that "the average boy" on movie-nights is much more active in the first hour, and slightly less active in the 7th and 8th hours of the five movie nights than in the corresponding hours of 35 normal nights. In other words, on the movie-nights, the sleeper postpones his settling down to rest, and rests only a little longer at a time in the later hours than he usually does. But Garvey and also Johnson and Weigand found that this is a perfectly normal occurrence. From the tables of the present authors, pp. 126 ff., it appears that the weighted averages of the averages of the individual subjects' active minutes per hour are subject to relatively large standard deviations, as are the corresponding averages of the individual subjects. Could one then choose five normal nights, which stand in the same temporal order as these five movie-nights, the records of which (by the Chi-squared criterion) would differ from the normal as widely as do the records of these five movie nights? Before we can test the

authors' appraisal of the influence of the movies, we need to see a just account of negative instances; and this remains to be presented.

The authors give (p. 153) some contradictory interpretations of the results. In Conclusion (1) they say that the movies were followed by changes in motility that cannot be attributed to normal variations, or to the "holiday-effect." This conclusion contradicts the data presented. In Conclusion (2) they say that a *decrease* (as well as an increase) of motility from the normal *diminishes* the "recuperative effects" of sleep; in Conclusion (3) they say that *decreases* in motility *enhance*, while *increases lessen*, these "recuperative effects." At least one of these assertions, must be false. Again, on p. 99 they say, and on p. 154 (Conclusion 5) they imply, that the greatest "recuperation" occurs in those hours in which the normal subject is usually the least motile; namely, his first two or three hours in bed. These assertions are not logically supported by any valid tests of "the recuperative effects of sleep" that are cited by the authors or known to me. Indeed, I have not yet seen exhibited a set of sensible operations that could define such a test.

Thus, certain parts of the verbal description contradict the authors' statistical exhibit. The statistics themselves contain some incongruities. Figs. 12 f., pp. 80 f. rank the seasons in different orders according to the average motilities of the different classes of subjects, although Fig. 12 purports to have been derived from Table 13. Also, the extrapolations in Fig. 12 contradict some facts. Again, Tables 16 f., pp. 100 f., are mislabelled. Their abscissae should read [(M/N)—1]%, not "M/N%." Fig. 28, p. 150, showing the relation between the mean length of the quiet periods to the number of active minutes per hour, does not show a restriction of the field of the equation given on p. 149. This restriction prevents the equation from asserting that if a child should lie still as long as an hour he would never stir again. But it also limits the field of the equation to intervals of exactly one hour, each of which ends in an active minute. But a generalized formula exists, which gives the same values as this one gives within the same field, which does not presuppose that a sleeper cannot lie still as long as one hour, and which is applicable to any temporal range that ends with the beginning of an active interval, whatever the unit-interval may be. It is the formula of Johnson and Weigand, cleared of the assumption that half of each unit-interval in which movement occurs is passive. Were the *original* formula of these authors, and the formula of the present authors, applied to the same records, they would yield two estimates of the mean rest-period which would differ by exactly half of one unit-interval: in this instance, by half a minute. The present authors say that the use of a 5-min. unit-interval instead of a unit-interval of one minute, "distorts" the exhibit. It remains to be shown how great this distortion is.

In summary, there is no doubt that the authors have collected a set of facts that have great potential value. Their presentation, however, is faulty. The comparison of average motilities among groups of different age, and in general, between any two individuals who differ greatly in either weight or stature is invalid. The influences of systematic variables on the motility of the same sleeper can be determined from the records, but the authors have used a set of statistical forms that do not exhibit the differences well. The historical sections of the book are especially poor. The historian discusses certain studies with which he seems to be unacquainted, since his account contradicts the original report. He sometimes quotes with approval the remarks of certain other historians who depend on secondary accounts

written by authors who have not read the original reports and who accordingly misrepresent them.

American University H. M. JOHNSON

Behavior Mechanisms in Monkeys. By HEINRICH KLÜVER. Chicago, Univ. Chicago Press, 1933. Pp. xvii, 387.

Phrased in non technical terms, the central problem of this investigation on "equivalent" stimuli is that of determining the nature and extent of changes one may make in external stimulus-situations to which monkeys have been making characteristic responses without altering the mode or the consistency of these responses.

Use was made of 16 animals, consisting of one lemur, and seven Java, four Cebus, two spider, and two squirrel monkeys. Their exact ages and previous experience were not known. Some were sexually mature when first employed and others became mature before the experiments were completed. Neither age nor sexual development appeared to be important factors in these experiments, but the rôle of previous experience could not be properly evaluated. Due observance of this point is made by the author in evaluating his results.

An idea of the author's method, called the "method of equivalent stimuli," may be obtained from the following illustration. A monkey is trained to pull to his experimental cage, from a distance of approximately 2 M., the heavier of two food boxes which in external appearance appear to be identical. While learning the amount of "pull" required by each box he hits upon the technique of tentatively pulling first one then the other string attached to the boxes and by so doing informing himself as to which is the heavier box. When the monkey can make the discriminative response correctly and consistently, he is given a series of crucial tests. In these the weight of one of the boxes is altered; if, for instance, the pulls required to move the two boxes in the training series had the ratio of 450:150 gm., those of the first crucial series may have the relation of 1350:450 gm. and those of the next, the relation of 150:50 gm., and so on. The ultimate object of these crucial tests is to determine the kinds and ranges of stimulus-constellations which, because of the unaltered mode of the animal's response, may be regarded as *equivalent* or *non-equivalent.*

Some of the most important procedures developed in the investigation are:

(a) The pull-in technique applied to movable stimulus-boxes, the differential responses being made to weight relata, auditory relata, and visual relata, (the latter having variations in area, color, brightness, figure-ground, and constancy versus intermittence).

(b) The pull-in technique with stationary stimulus-boxes (the food-box alone being movable) requiring differential responses to visual relata that at all times remain in a fixed position. Although little used, this method was deemed particularly suitable for the study of complex visual relata that can be presented only with elaborate, non-portable instruments.

(c) The pull-to-stop technique, in which the food-box is pulled in until a sudden change takes place, such for instance as a constant light becoming intermittent. Trials in which the appearance of the box remains constant and trials in which some variation is suddenly introduced follow each other in chance order. As soon as the pull-to-stop reaction to the change is established, a variety of crucial

tests are given to discover equivalent and non-equivalent stimulus-constellations. Since only *one* box is used the method has the advantage of dispensing with a number of troublesome control experiments necessary when two or more boxes are used to check up on possible secondary cues.

(d) Pull-in techniques in which there are several stimulus-boxes from which the correct one must be selected. The technique, called one for presenting "strata stimuli," vaguely reminds one of multiple-choice experiments. But the ideas prompting the experiment as well as those expressed in the interpretation of data collected, show that the author has gone far beyond the traditional psychologizing issuing from previous multiple-choice experiments.

(e) Methods for measuring time relations in the monkey's successive comparisons while using the pull-in techniques; for determining visual and tactual acuity for small objects, such as threads; and for evaluating the animal's ability to utilize objects as tools.

In the brief but masterly discussion of principles underlying the interpretation of animal data, Klüver displays a scholarly grasp of current and traditional fields of psychology. In this respect, the monograph is probably unsurpassed by any monograph heretofore presented in the field of animal psychology. The discussion of the sensory properties of stimulus-situations reveals the difficulties one encounters in defining the animal's perceptual world; *i.e.* the *Umwelt* of Von Uexkull. That on reactions to abstract relationships which persist in stimulus-situations, the physical properties of which have been radically changed, gives evidence of thorough grounding in the contributions of Köhler, Koffka, Révéz, and others who have influenced recent thinking on this and related subjects. The problems of "relative" and "absolute" reactions, "approximate constancy" in perception of objects, "generalization" and "differentiation," "field" properties of sensory stimuli, and the more general problem of "intelligence" are searchingly examined in relation to his collected data. In generalizations and summarizations of what may be taken as evidence for or against a given concept or theory, the author is ultra-conservative.

One noteworthy omission in the author's discussion is the application of his results to the old problem of "the transfer of training" which has already been extensively investigated by several students of animal psychology. Many of his data shed light upon this venerable problem, in one or more of its aspects, and conversely some of these studies on transfer anticipate and illuminate the results obtained by Klüver.

Fortunately, we may expect further contributions on these same experiments since the animals were carried along and retested from time to time and eventually were thoroughly re-examined after undergoing brain operations. As pointed out by Dr. K. S. Lashley, who wrote the introduction for the monograph, "The most immediate value of the study is in laying a foundation for investigations of the neurophysiology of behavior."

Stanford University CALVIN P. STONE

Some Parent-Child Relationships as Shown in Clinical Case Studies. By MARIAN J. FITZ-SIMMONS. New York, Teachers College, Contrib. Educ., No. 643, 1935. Pp. xii, 162.

The investigation here described had a two-fold purpose, first, to develop an objective device by means of which the information included in the case records of

children referred to child guidance clinics could be classified with reference to the attitudes of the parents toward the child; and secondly by utilizing the scale thus devised in the study of a selected group of cases, to ascertain the effect of such parental attitudes upon child conduct.

Following the usual review of the literature, there is an interesting discussion of the case history as a tool for scientific research. Here the writer seems to have been torn between the desire to prove the validity of her data and the recognition of certain inherent sources of error. She correctly points out that certain of these errors are not essentially different from those encountered in the administration of intelligence tests but fails to note that similarity in kind does not necessarily indicate equality of degree. An even more serious fallacy is to be found in her tendency to confuse the case record as it commonly is with the case record as it ideally might be. Thus, on page 29 she notes that "if there were a way of determining numerically the reliability of case records, inconsistency due to response errors in the examinee would most certainly be a cause of statistical unreliability but would not invalidate the record for research because that very inconsistency in the examinee may be an important contribution to the knowledge of his personality traits." With this statement I am in hearty agreement, but the conclusion drawn in the sentence immediately following is true only in a very limited sense. The author states that "Because case material collected by two different workers differs in slight or large measure, it is not safe to draw the conclusion that case material is not reliable in the sense of being dependable or trustworthy as a picture of the true condition of the examinee. Errors noted in conflicting statements made to different people might produce statistical unreliability but *do not reduce the validity of the record.*" (Author's italics.) The change from the subjunctive to the unconditional declarative statement is quite unwarranted. Unless two or more independent case histories are actually obtained so that such inconsistencies can be discovered, and until research necessary for the interpretation of the clinical significance of such inconsistencies has been carried out, their presence in a clinical record must inevitably reduce the validity of that record. Since neither of these requirements is fulfilled in the usual study employing such records (the present one is no exception to this rule), the conclusion cited can hardly be accepted as valid for conditions as they now exist.

In Chapters 3 and 4, a very complete description of the construction of a "Guide for the estimation of parents' attitudes" is presented. The attitudes considered are those having to do with parental affection—the so-called attitudes of acceptance, rejection or over-protection of the child. Over-protection is assumed to be a compensatory mechanism by which a real antagonistic attitude is concealed. It was found that when two workers with similar training evaluated the data on parental attitudes included in 48 case records, according to the method outlined in the Guide, their estimates agreed to the extent represented by a contingency coefficient of $+0.80$. This, of course, is not an indication of the reliability of the method in terms of the factual content of the records; it corresponds rather to the amount of agreement between two persons in scoring the same test blank. It does, however, indicate that the Guide provides a fairly objective way of evaluating the records as they stand.

The second part of the study presents the results obtained by scoring the records of a selected group of 94 children studied at three Child Guidance Clinics. Only

"full study" cases of children in good health with IQs of 85 or higher and coming from unbroken homes in comfortable or affluent financial circumstances were included. There is an excellent analysis of the extent to which the conditions of selection differentiated the selected cases from the general clinic population. The attitudes of the parents as estimated by means of the guide are then compared with the reported behavior of the children. The results are disappointingly meager. In the first place, it was found that in only 64 of the 94 cases originally selected was there sufficient data to permit a classification of the parents' attitudes. These cases were subdivided into 14 classes representing different combinations of parental attitudes. Only three of these sub-groups were judged to include a sufficient number of cases to warrant comparisons. In 13 cases both parents were judged to have a negative (rejective) attitude toward the child; in 9 cases the mother was judged to be over-protective, the father negative; and in 12 cases the mother's attitude is classed as negative, the father's positive. Children belonging to the first group appeared to lead in the average number of reported forms of "aggressive" behavior such as lying, stealing and cruelty; children belonging to the second group showed a greater number of reports of "withdrawing behavior"; and children belonging to the third group slightly more undesirable habits such as thumb sucking and nail biting. None of the differences is large, however, and in view of the small number of cases in each group and the absence of any evidence as to the seriousness of the behavior in question (all cases of stealing, for example, being classed together, regardless of the frequency or seriousness of the offenses), little confidence can be placed in the findings.

The study is an excellent example of the many futile attempts that are being made at present to utilize case records for statistical purposes without first insuring that the records shall be taken in such a way as to render them amenable to statistical treatment. Another serious criticism that may be made of this type of approach is the absence of valid control data. In the present study the two comparisons that would be of greatest interest are lacking: (1) the attitudes of parents of "problem" children as contrasted with the attitudes of those whose children appear well-adjusted; and (2) the conduct of children whose parents are judged to be negative in their attitudes as compared with that of the children of parents with a positive attitude. The author is to be congratulated on having chosen a problem that unquestionably is important from the clinical standpoint; but it is unfortunate that a more suitable approach to its solution was not selected.

University of Minnesota FLORENCE L. GOODENOUGH

The Frontiers of Psychology. By WILLIAM MCDOUGALL. New York, D. Appleton-Century Co., 1934. Pp. xi, 235.

Psychology according to Professor McDougall has two distinct frontiers, one which faces the sciences of Nature: the biological sciences, the inorganic sciences, mathematics, etc.; and the other the sciences of mind: the social sciences, anthropology, history, etc.

On the frontier facing the natural sciences, the psychologist (McDougall) encounters such problems as mathematical prodigies, the astronomer's universe as a construction of his mind, space and time, force and energy, and similar physical abstractions. Some of the pests which disturb Professor McDougall's exploration on

this frontier include Needham, whose mechanistic biology seems to have dispensed with truth altogether, substituting for it a convenient fiction which has for its justification merely a mental twist or scientific prejudice; and Jeans, whose simple-minded psychology leads him to speak of "purely mental" concepts, as if there could possibly be any other kind. Before judging Jeans too harshly, however, should we not make sure that the psychologist (McDougall) is not equally naïve about his physics?

Professor McDougall writes: "Energy is (quantity of) ability to perform work. All energy is tending to conversion into the form of energy called entropy, according to the inescapable second law. Eventually, all energy must take this form. The quantity of energy, having been strictly conserved, will then be still the same; but all performance of work will be impossible; there will be in the universe no ability to perform work. That is to say the universe will contain the same immense quantity of ability to perform work as at the present moment; but it will at the same time contain no ability to perform work" (p. 140). In the first place, entropy is not a form of energy, according to the physicist. Secondly, there is no contradiction of terms here since the second law of thermodynamics states that the energy is simply less available to perform mechanical work under the conditions outlined here. Finally, before anyone quarrels with the physicist on this ground, should he not be thoroughly familiar with the relative nature of the concepts of energy as the physicist understands them?

On the frontier facing the mental sciences the psychologist (McDougall) discovers the problems of value and progress. Progress came about through a maturation of creative thinking, which in turn is a characteristic now inherited by northern Europeans through some thirty generations of adaptive effort to acquire it. Evidence for all this is found in Professor McDougall's own experiments on rats. Although these experiments have not yet been corroborated and have been severely criticized, Professor McDougall writes: "I submit that we have now such positive experimental evidence in support of the Lamarckian theory as justifies us in applying it to the interpretation of the problems of history and sociology" (p. 210). To which a deserving reply seems to be: "I submit that we have not."

Psychologists who believe with Needham that science is "essentially mathematical, mechanical, deterministic, quantitative, abstract," are likely to agree that the most disturbing pest on the mental frontier is the explorer himself. They will feel that a happier choice than Professor McDougall could be made to represent psychology in a gathering of scientists that includes such eminent names as Boas, Eddington, Jeans, Needham, Pavlov, Planck. Their choice would not include Jung or Freud, two associates cited with approval by our present explorer.

University of New Mexico GEORGE M. PETERSON

For Stutterers. By SMILEY BLANTON and MARGARET GRAY BLANTON. New York, D. Appleton-Century Co., 1936. Pp. xi, 191.

For Stutterers is directed primarily to those who suffer from the handicap named in the title, but there is reason to believe that it will be read and used for reference by many psychologists who have no contact with speech-clinics. The reason for this belief is to be found in the Blantons' thesis that stuttering must be regarded as a symptom of psychological disorder rather than as the disorder

itself. To introduce the reader to the nature and the causes of these underlying disorders, the authors have found it necessary to write what constitutes actually a brief manual of psychological hygiene. In this latter guise the book is likely to find a sympathetic reception among psychologists who have no connection with the speech-clinic.

The Blantons' theory of the cause of stuttering is simply stated. "All of the parts of the body that are utilized by speech are only borrowed for speech and have their first and most vital use in some other field of activity. . . . Speech only borrows these areas, but surrenders them at once, in case of emergency, to their primary function." "We believe that the physical symptoms of stuttering are caused by an emotion that blocks off this control of the cortex over the thalamus and lower nerve centers and allows the primitive patterns which have been repressed to come forth and interfere with the movements of speech."

If the reader grants the reasonableness of this explanation, he is *ipso facto* compelled to grant that "it is futile to 'treat' the speech and futile to give any treatment that does not have as its aim the elimination of the underlying emotional difficulty." The authors propose two modes of treatment: (1) psychoanalysis; and (2) individual guidance. Psychoanalysis, which must be carried only by a qualified analyst, "is not indicated for all stutterers by any means, nor is it necessary for all." For most stutterers, individual guidance at the hands of a physician of adequate training is described as the most suitable method; it involves a conjoint search by physician and patient in an effort to discover the sources of emotional distress which underlie the stutter. Other theories of the causation and treatment of stuttering are discussed and evaluated, and the stutterer is repeatedly warned of the dangers of interference by well-intentioned but untrained amateurs.

Along these general lines the authors have written a conservative and coherent résumé of present-day knowledge of the etiology and therapeutics of stuttering. Their rich clinical background is apparent in the chapters on the development of conscience, on anxiety, and on various problems of emotional adjustment in the family and in society. The closing chapters are devoted to a consideration of what can be done by the stutterer himself, by his parents, and by his teachers in the alleviation of this distressing symptom.

The general tenor of the book is marked by a pleasing absence of dogmatism. Such chapter-headings as "What We Believe to be the Cause of Stuttering" show that the authors are content to leave theory as theory and not to parade it as established fact. One gains the impression that the authors have weighed existing theories carefully and have proposed their own in the light of existing evidence; their good faith is further indicated by their confession that no one theory has yet been formulated to cover all the known phenomena of stuttering. The terminology employed is that of the psychoanalyst, but it is remarkably free from the multiplication of causal agencies so common in psychoanalytic writing. Although it is sufficiently elementary for the average layman, the book has definitely not been 'written down' to any group of readers. On the contrary, it reflects the broad training and extensive clinical experience of the authors in a fashion which will make it an important addition to the bookshelf of the psychologist who deals with problems of psychological hygiene.

Cornell University JOHN G. JENKINS

Stammering and Allied Disorders. By C. S. BLUEMEL. New York, Macmillan Co., 1935. Pp. vi, 182.

In 1913 Bluemel wrote that stuttering resulted from transient auditory amnesia. In 1930 he attempted to clarify this theory by explaining that the momentary amnesia was due to a recoil of the thought from the mind, and not necessarily due to a natural weakness of the memory for sound.

In the present book he attempts to clarify the theory still further by a change in terminology, introducing the conditioned reflex term, inhibition. His present theory is that stuttering is due to a physiological collision between the conditioned reflex initiating speech and inhibition which impedes it.

The first part of the book deals with the conditioned reflex in relation to language and speech development, inhibition, and the conflict between conditioned reflex and inhibition. Then follows a discussion of primary and secondary stuttering and the clinical aspects of inhibition. Bluemel concludes his work with a discussion of various prevalent theories of stuttering and the treatment of this defect according to the conditioned reflex theory.

Bluemel presents his concept of stuttering clearly and logically. There is not, however, one shred of experimental proof offered in support of it. Neither is there any systematic evidence offered that the practical application of the theory alleviates stuttering conditions. The book represents mainly another addition to the already large mass of unsupported theorizing in regard to the nature and management of stuttering.

University of Iowa LEE EDWARD TRAVIS

Developmental Psychology: An Introduction to the Study of Human Behavior. By FLORENCE L. GOODENOUGH. New York, D. Appleton-Century Co., 1934. Pp. xvii, 619.

This is a "first book in psychology." Its novel feature is that it attempts to present the subject matter of psychology genetically. Of the 27 chapters, 19 take the child from prenatality to adolescence. Then follow chapters headed "Educational and Vocational Guidance," "Motivation of Behavior at the College Level," "Adult Behavior and Social Customs," "The Maturation and Decline of Abilities," "Mental Disease and Old Age."

One didactic feature of the arrangement is commendable; every chapter starts by raising a number of questions or problems. But in the section "To the Student" the author states that "Psychology, as it is taken up in this book, is especially concerned with finding out the reasons for human action." Further "it can and does tell us much about the principles to which human behavior conforms, and from a knowledge of these principles we can at least infer, in most cases, why the action in question might reasonably have taken the form it did" (p. viii).

Where and how is the beginner in psychology to get hold of these magic principles? "So as you read, ask yourself other questions beside those given at the beginning of the chapter. 'Have I ever known people to act like this?' 'Does this theory or this principle agree with what I have observed?' Find out what other people think or know about the matter. Discuss. Argue. Better still consult additional authorities; go to the library and find out what these authorities have to say. That is the way to make learning real."

The power of the laboratory and the rigor of scientific method in resolving important problems is not mentioned. The beginner is urged to try logic, debate, argumentation; or to consult "other authorities." As one reads section after section he is continually impressed by the number of important investigators and their contributions to which no reference is made. How, for example, does a child construct his world of space, form, movement? "Perception is built up from experience" (p. 272) but how the welter of "experiences" get themselves systematically organized is not disclosed. In fact the dogmatic and positive form of statements of this sort is certain to clash with the advice to the student to find out what other real authorities have to say about such basic problems. On p. 241 this generalization is printed: "You and I are what we are today—see as we do, think as we do, behave as we do—because we grew the way we did."

The account of the manner of establishing any complex coördination is interesting and revealing. On p. 372 we read that "the overlapping of elements in a pattern reaction . . . is one of the chief reasons why a well learned action seems to be done so smoothly and evenly as contrasted with the uneven, jerky, disconnected movements that are characteristic of the beginning stages of learning." "The secret of a smooth and well organized motor performance lies in the ability to keep the attention centered on what is coming rather than what is being done at the moment."

The range of topics covered is so large that it is naturally impossible to treat each one exhaustively, but the omission of important facts, the challenging of the student's best powers to look at the problems of genetic development from all angles (the treatment, for example, of how older children learn) are things which could and should be done better in a beginner's textbook. The style is clear and many illustrations are used so that the book could be recommended to a none too critical person who wants to read a book on child psychology. As a textbook for college students, particularly as their first book in psychology, no.

Ohio State University S. RENSHAW

The Range of Human Capacities. By DAVID WECHSLER. Baltimore, Williams & Wilkins Co., 1935. Pp. ix, 159. ,

Taking the ratio of the 2nd and 999th of every 1000 cases observed as the best means of expressing the total range of human capacities, Wechsler, by analyzing studies of some 89 traits, finds that the ratios between extremes have striking recurrent values which, on the average, fall roughly within the range 1.2:1 to 2.5:1. He believes that the range of human capacities is exceedingly small, and that there are calculable limits to human variability which are biologically determined and partake of the characteristics of natural constants. Without questioning the validity of the technique, one difficulty in evaluating these findings lies in the limited number of studies (7 in all) of perceptual and intellectual abilities presented in the analysis, Wechsler having omitted a considerable number as inadequate for his purposes. The paucity of adequate studies is, no doubt, in part due to the lack of satisfactory units of measurement for many mental abilities, and Wechsler suggests the method of psychophysical correlation as a solution for this dilemma.

Another important thesis of the author is that the age curves of such mental abilities as have been measured indicate that intellectual capacity begins to decline earlier rather than later than most physical capacities, and there is no evidence

whatsoever for the belief that the average man maintains his intellectual or physical vigor to the end of his natural life. In short, "the acme of a man's intellectual and artistic virility . . . would certainly not be much over 40, and probably much nearer to 35" (p. 93). Wechsler here reasons from Galton's data, the original Army Alpha results, and more recent studies such as Miles and Jones. He also argues that the greatest contributions of outstanding members of society are made much earlier than is usually stated.

The study of the maturation and decline of mental and physical abilities is today deservedly receiving more attention. Wechsler's work indicates a woeful lack of competent data in this field. This is further exemplified by present day conflicting opinion. Does "life begin at forty," or is the individual well on to decline by then as Wechsler contends? The unequivocal test of his conclusions cannot be made at present, but must await the judgment of more, competent, data than there are to be found in the journals at present.

Well written and carefully developed, even for those whose knowledge of statistics is limited, this volume is indeed a contribution to the field. The tabular material is well coördinated in 3 appendices which do not conflict with the organization of the material in the text. The fundamentals of probability, "normal" curves, and other basic concepts are so well handled as to make the book admirably suited for collateral reading, particularly for courses in individual differences.

Brooklyn College EDWARD GIRDEN

The Content of Motion Pictures. By EDGAR DALE. New York, Macmillan Co., 1935. Pp. xvi, 234.

Children's Attendance at Motion Pictures. By EDGAR DALE. New York, Macmillan Co., 1935. Pp. x, 81.

Two further studies carried out under the auspices of the Payne fund for the study of the effect of motion pictures upon American youth are here presented within the covers of a single volume. As is indicated by the title, the first is concerned with an analysis of the content of motion pictures. Three different techniques were employed. First, by means of the summarized descriptions included in an abstracting service for motion-picture exhibitors, an analysis was made of the changing trends in the major themes of feature pictures over a period of years. For this purpose, a sampling of 500 pictures for each of the three years, 1920, 1925, and 1930 was used. The proportion of pictures in each sample falling within each of a number of defined types showed only small changes over the 10-yr. period. In each case, "romantic love" was the most frequent major theme but the proportion of the total sample classified under this heading dropped from 44.6% in 1920 to 32.8% in 1925, and 29.6% in 1930. "Crime pictures" held second place and "sex pictures" third in all three years, with proportions of about 25% and 15% respectively. "Comedy pictures" held fourth place as feature pictures and the proportion of comedy pictures increased slightly during the period considered. These four themes accounted for more than 90% of the entire sample studied.

A more intensive study was made of 115 pictures that were actually viewed by from one to three observers, and a still more intensive analysis was made of 40 of these for which dialog scripts were obtained from the producers. Such factors as locale of the story, economic status, age, and clothing of the characters, point of view represented toward various social problems, techniques of love-making,

etc., are enumerated in detail. From all this the author comes to the conclusion that the typical movie depicts a fairy-tale sort of life that has small relationship to real existence. Its chief appeal is to the "escape" motive. Although the author recognizes that this may be a legitimate function up to a certain point, he is inclined to think that the situation may have passed beyond the potential danger line.

The data for the study on children's attendance at moving pictures consist of questionary reports on attendance that were collected from over 50,000 Ohio school children and tabulations of the population attending certain theaters in Columbus. The author concludes that children constitute a slightly greater proportion of the moving-picture audience than would be expected on the basis of their frequency in the general population. Boys attend somewhat more frequently than girls and they are more often unaccompanied. It is unfortunate that this study is based only upon an Ohio population, since state laws regarding the admittance of children to moving-picture theaters vary widely.

Together, the two studies raise a number of important social and educational questions. Certainly, the composite photograph of the moving picture of 1920-30 as presented by Dale does not commend itself as material suitable for presentation at weekly intervals to children whose ideas and ideals are still in process of formation.

University of Minnesota FLORENCE L. GOODENOUGH

Motor Adaptation and Accuracy: An Experimental Study. By HARRY MCNEILL. Louvain, L'Institut Superieur de Philosophie, 1934. Pp. xiii, 303.

This volume is the second in a projected series of studies of bodily movements under the direction of Albert Michotte at the University of Louvain. It aims to discover the general relationships existing between the different types of spatial and temporal fields in which an object to be hit (or target) and the human subject are placed.

The study is divided into two parts. The first part analyzes the movements and errors made in attempting to hit targets or spots placed on a tabletop and arranged in different physical and temporal patterns. Means were provided for recording the trajectories of movements and the location of hits. It was found that the actual movement is always a compromise between following the indicated pathway, as perceived, and the simplest combination of bodily movements which will approximate this pathway. Diminishing the clearness of perception through limiting the time for observation, dimming or distorting vision, fatigue, hurry, inattention lack of interest, etc., interferes with the accuracy of the movements and tends to change the pathway "in the direction of the limit of motor simplification." If the model is modified to conform to the motor approximation of it there is further simplification in the direction of yet more generalized mass movements. If, on the other hand, the perceived pattern of the model is varied in the opposite direction the approximation of the original pattern is improved and errors corrected. A method of discovering the trend toward simplification is proposed through increasing imperception or "blind" movement. When a rhythm in movement is imposed the length of the segments is varied so as to achieve uniform speed.

Errors in hits are seen as either systematic or fortuitous. The former are due to, and in the direction of the constant force making for simplification of movement. The latter are ascribed to chance fluctuations in motor-set and mechanical interfer-

ences. Factor analysis of data should apparently be of value in testing this hypothesis.

The second part of the study deals with hitting at keys, somewhat as in typewriting but on a larger scale. While analysis in this case is less complete the indications are approximately the same as for the target experiments. The finding that increased amplitude of movement increases error systematically seems of obvious practical significance for typewriting, even though the study avowedly avoids excursions into the field of applied psychology.

Increased familiarity with the approach to motion study at Louvain breeds inevitably increased appreciation. Greater insight into behavior can come only through painstaking studies of this type. While the treatment of details in this case seems rather ponderous, and at times rather obvious, it is more acceptable than nonchalant paucity. Verbal generosity in this case, however, falls short of an adequate review of related studies at other institutions. A more thorough statistical analysis of results should also add much to the value of the study.

University of Colorado T. H. HOWELLS

You Must Relax. By EDMUND JACOBSON. New York, McGraw-Hill Book Co., 1934. Pp. xv, 201.

You Must Relax is a successful abridgement and popularization of the thesis presented in the same author's technical volume *Progressive Relaxation.* The former was obviously written for Jacobson's colleagues in medical and academic circles; the newer edition is written--under some protest—in response to the demand for a simple presentation comprehensible to the man in the street. For him the author depicts the familiar symptoms of muscular hypertension and describes in adequate detail the attempt to alleviate these symptoms by means of systematic instruction in the art of muscular relaxation. Pointing out that mere verbal commands to rest or to relax may fail utterly to relieve muscular hypertension, Jacobson outlines a series of training programs for those who wish to learn to relax. He cautions the reader that these programs will require time and warns him that they are best carried out under the guidance of a trained physician.

It must be realized—and on this point Jacobson is insistent—that the process is *not* one of suggestion. The patient is trained first to recognize "sensations of tension" in a particular muscle or muscle group, and is then instructed in relaxing by securing a diminution or absence of these "sensations" in that particular muscle-group. Only after the patient has mastered the relaxation of that group is he permitted to undertake the task of learning to relax some other localized group of muscles. Photographs indicate the location of the musculature involved in the movement of the several bodily members and depict the initial exercise by which the patient is trained to recognize the tensions involved. Once training in this *progressive relaxation* has gone far enough to permit general relaxation in the prone position, the patient is provided with a course of study in *differential relaxation* to be applied while sitting or standing.

The later chapters of the book are concerned with specific cases in which relaxation has proved its value in relieving patients who were suffering from the results of excessive worry, insomnia, and "nervousness," as well as those afflicted with spasticity of esophagus or colon or with high blood pressure. One chapter serves to introduce the reader to Jacobson's technique of recording electrically the action-current patterns by means of fine electrodes inserted directly into the muscle in-

volved. He is informed of the correlation between recorded frequency of discharge and intensity of reported tension, as well as of the relationship between psychological activity and tension in the muscles of the limbs and in those involved in speech. The literary style of the volume is clear, coherent, and simple enough to permit any layman to grasp the author's thesis and to follow his delineation of a definite program of training. Jacobson's continual insistence upon limiting the method to muscular relaxation and his cautions regarding premature enthusiasm for early results should do much to safeguard this promising method from destructive inroads by chronic hypochondriacs in search of the latest therapeutic miracle.

Cornell University JOHN G. JENKINS

Genealogy of Sex. By CURT THESING. Translated from the German by Eden and Cedar Paul, with an introduction by Smith Ely Jelliffe. New York, Emerson Books, 1934. Pp. xii, 286.

Dr. Thesing's book on the *Genealogy of Sex* deals solely with the reproductive aspect of animal behavior. It gives us, as Dr. Jelliffe so aptly expressed it in his introduction, a "biologic panorama" of the methods by means of which life, from protozoa to man, "tends to perpetuate itself."

The first chapter is devoted to an exposition of the methods of asexual and sexual reproduction, and the second, to a description of the germ cells. In the third chapter, Dr. Thesing traces the genesis of the sexual impulse through the hermaphroditic and intermediate stages to the sexually distinct. With these chapters serving as an introduction, the author then turns to his problem, to which he devotes the remainder of the book, the description of the sex mechanisms in the various forms of animal life. Specific instances are chosen from the various animal phyla, genera, and species to mark the mile stones in the development of these mechanisms. The last chapter of the book is anthropological and deals exclusively with primitive man.

The book is interestingly written and profusely illustrated by drawings and by photographs of animals in sexual play and embrace. Simple and non-technical language is used as far as possible by the author (or translators) so it may be understood by the laity—for whom it was obviously intended. The scholar will miss the bibliographies and references which have been omitted, and his risibilities will frequently be aroused by reference to the copulation of animals as "weddings," and to the male animals as "husbands" and to the female as "spouses" and "wives."

K. M. D.

A Textbook of Clinical Neurology with an Introduction to the History of Neurology. By ISRAEL S. WECHSLER. Third edition reset. Philadelphia and London, W. B. Saunders Co., 1935. Pp. 826.

This new edition of a standard work in clinical neurology will be of especial interest to psychologists not only because of its general merits but because it contains an excellent brief chapter on psychometric tests written by a psychologist, Dr. David Wechsler. The fifth and last section of the book on the neuroses, as well as the appendix, which presents an introduction to the history of neurology, will also be of interest to many readers of this journal.

The whole and particular aim of the volume is *clinical.* This means that on page after page the laboratory neurophysiologist or physiological psychologist will

find statements which will disturb him, as one used to the controlled methods of pure science. For example, sensory tests such as the "watch test" or the "whispered-voice test" are recommended. The inaccuracy of such measures is the more surprising in that on following pages the chemical analyses of body fluids are given to the third decimal place.

The use of the term consciousness in this book will also be questioned by many psychologists. Consciousness is said to be "undoubtedly a function of the whole brain." In this connection other terms referring to conscious processes, such as sensation, are sometimes employed where most neurophysiologists would use the term nervous impulse.

As an indication of the rapidity of change in neurology, it is interesting to notice that even in this 1935 book no reference is given to Fulton's work on the syndrome of the premotor cortex or to the controversy which it has aroused. Similarly, in the discussion of the optic system, cortical projection fibers are shown without question as arising in the thalamus proper as well as in the lateral geniculate bodies.

The volume summarizes a large mass of clinical and observational material concerning the relationship between the structure of the living body and normal and abnormal behavior. Much of this material is directly relevant to behavioral psychology, and, in spite of the fact that because of its clinical nature it cannot be given experimental verification, it surely should not be neglected by any student of human mental life.

Brown University LEONARD CARMICHAEL

Wahrnehmung und Gegenstandswelt: Grundlegung einer Psychologie vom Gegenstand her. By EGON BRUNSWIK. Leipzig, Deuticke, 1934. Pp. xi, 244.

Perception of an object (*Gegenstand*) has the advantage of being quick and comprehensive, but also the disadvantage of being variable from time to time. Its measurement, on the other hand, is not so variable. The classical psychophysics was mainly devoted to discovering the amount of precision in perception (*Wahrnehmung*) when used in lieu of measuring (p. 7). The author's perception-psychology (*Psychologie vom Gegenstand her*) aims at furnishing an inventory of perceptual efficiency (p. 42). In estimating the seen relative size of objects which are at different distances from us, our precision is rather high, for educated adults about 96% of the true size; but in sounds it is only about 50% (p. 70). The author refers to the imagined object used as intermediary in such cases as "Z-value" (*Zwischenwert*). He says that "psychology is the only science which needs this concept of Z" (p. 82). "What is my immediate consciousness when looking at a clock—the time or the placement of the hands?" (p. 90). The foregoing is one of his examples of Z-values. "Perception can no more be reasoned with than can a circuslion. Even when we actually see that someone puts his finger on a weight held in our hand, the weight at that moment seems to become heavier" (pp. 118 f.). The author wants to substitute for the *Gestaltpsychologie* a system of thought which is much broader, older and more promising, and which he calls "Psychology with a proper recognition of *all* that is in the object," a psychology which asserts that *the whole* is *not always* perceived unanalyzed. The book is to be recommended to all those *Gestalters* for whom one may still hope that they broaden out.

University of Missouri MAX F. MEYER